



**University of  
Nottingham**

UK | CHINA | MALAYSIA

**High throughput surface mass spectrometry-based proteomics  
& metabolomics for biological applications**

**Joris Meurs, BSc., MSc.**

Thesis submitted to the University of Nottingham for the degree of Doctor of  
Philosophy

17th June 2021

# Acknowledgements

At first, I would like to thank my supervisors Dr Dong-Hyun Kim, Professor Dave Barrett and Professor Morgan Alexander for their help and support to make this a successful project. Further special thanks to Dr Catherine Ortori and Dr Salah Adbelrazig for their technical support on the Nanomate and the Q Exactive. Without them there would not have been any analysis done! Also, a big thank you goes to Dr Jordan Thorpe, Dr Aishah Nasir and Dr Sara Pijuan-Galitó for supplying the cell data and for fruitful discussions to get the most out of it. Another shout out goes to Dr Laurence Burroughs for the help on the contact printing, otherwise there would have been any arrays available. Speaking about arrays, I am grateful to Dr Pavel Levkin and his group to initially supply us with some free Droplet Microarrays for testing, which eventually led to the first paper of my PhD! Of course there is also Dr David Scurr to thank for his help on the ToF-SIMS and 3D OrbiSIMS, from which also a nice paper come forward eventually! Furthermore, there all the administrative/technical people (Liz, Sian, Graham, Paul, Amy & Julia) to thank for sorting out many important things around lab stuff, conference and all kinds of paper work! Finally, I would like to thank my lovely wife Katarzyna for all her support during my PhD. And there are also our pug Benia and daugther Raila who bring a lot of joy and happiness in our life!

# List of abbreviations

2-DMPA	2,2-dimethoxy-2-phenylacetophenone
3D OrbiSIMS	Orbitrap secondary ion mass spectrometry
ACN	Acetonitrile
AGC	Automatic gain control
AmBic	Ammonium bicarbonate
AUC	Area under the curve
Cyt c	Cytochrome c
DA	Discriminant analysis
DAPI	4',6-diamidino-2-phenylindole
DDA	Data-dependent acquisition
DESI	Desorption electrospray ionisation
DIA	Data-independent acquisition
DMEM/F12	Dulbecco's modified eagle medium/Nutrient mixture F12
DMF	<i>N,N</i> -Dimethylformamide
DMSO	Dimethylsulphoxide
DTT	Dithiotreitol
ECM	Extracellular matrix
ESI	Electrospray ionisation
FA	Formic acid
FBS	Fetal bovine serum
FDR	False discovery rate
FGF-2	Fibroblast growth factor 2

FT-ICR	Fourier transform ion cyclotron resonance
GC	Gas chromatography
GRAVY	Grand average of hydropathy
H <sub>2</sub> O	Water
IAA	Iodoacetamide
IPA	Isopropanol
hESC	Human embryonic stem cells
hiPSC	Human induced pluripotent stem cells
LC	Liquid chromatography
MALDI	Matrix-assisted laser desorption ionisation
MeOH	Methanol
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MSI	Mass spectrometry imaging
MW	Molecular weight
NaCl	Sodium chloride
PCA	Principal component analysis
PLS	Partial least squares
PSM	Peptide-to-spectrum match
ROC	Receiver operating characteristic
ROCK	Rho-associated, coiled-coil containing protein kinase
SDC	Sodium deoxycholate
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
TCA	Trichloroacetic acid
TFA	Trifluoroacetic acid
TGF- $\beta$ 1	Transforming growth factor $\beta$ 1
TIC	Total ion count
VIP	Variable importance in projection



# Constants

Constant	Value
$\pi$	3.14159265359
$h$	$6.62607004 \times 10^{-34} \text{ m}^2 \cdot \text{kg} \cdot \text{s}^{-1}$
$e$	$1.60217662 \times 10^{-19} \text{ C}$
$m_e$	$9.10938356 \times 10^{-31} \text{ kg}$

# List of essential amino acids

Amino acid	One-letter code
Alanine	A
Arginine	R
Asparagine	N
Asparatic acid	D
Cysteine	C
Glutamic acid	E
Glutamine	Q
Glycine	G
Histidine	H
Isoleucine	I
Leucine	L
Lysine	K
Methionine	M
Phenylalanine	F
Proline	P
Serine	S
Threonine	T
Tryptophan	W
Tyrosine	Y
Valine	V

# List of Figures

1.1	Schematic representation of a material microarray fabrication. (A) The material solution is collected from the source plate. (B) The solution is then dispensed onto the printing substrate. Adapted from Algahtani <i>et al.</i> [15]. . . . .	4
1.2	Schematic representation of different printing methods. In contact printing, the solution is released upon contact with the substrate. For piezoelectric printing, a voltage is applied on the piezoelectric pin which causes the pressure to increase within the solvent chamber permitting release (jetting) of picoliter droplet from the pin. The droplet volume can be tuned by altering the applied voltage. . . . .	6
1.3	Schematic representation of the principle of SIMS . . . . .	10
1.4	Jablonksi diagram electronic states used in Raman spectroscopy. Adapted from [44] . . . . .	11
1.5	Schematic representation of the principle behind XPS. Adapted from [49]	12
1.6	Schematic representation of the principle behind water contact angle measurements. Adapted from [50] . . . . .	13
1.7	Schematic representation of the potential cell-material response mechanism. (A) First, (a set of) proteins attach to the synthetic polymer surface and serve as (B) a scaffold for the cells to the attach to. . . . .	16

1.8	Approaches for MS-based protein analysis. (A) Top-down strategy. Intact proteins are introduced into the MS and dissociated via gas-phase collisions. The fragment ions represent a part of the amino acid sequence of the protein. (B) Bottom-up strategy. Intact proteins are first (enzymatically) digested into peptides which are then introduced into the MS and fragmented. The MS/MS spectrum can then be used to identify the peptide based on the amino acid sequence from which a protein identity can be inferred. Adapted from [81]	18
1.9	Nomenclature of fragment ions of peptide as defined by Roepstorff & Fohlman and Biemann. Adapted from [94]	20
1.10	Schematic representation of a quadrupole mass analyser. Adapted from [96]	21
1.11	Schematic representation of the Orbitrap mass spectrometer	22
1.12	Schematic representation of an electron multiplier. Adapted from [96]	23
1.13	Schematic representation of detection in a FT instrument	23
1.14	The .FASTA format explained	25
1.15	The use of TMT for relative quantification of proteins. (A) Example of a tandem mass tag. The amine-reactive group is used for covalently bind to the peptide. Each tag has a different reporter ion which make the tags distinguishable at MS <sup>2</sup> . The balance part is essential to make up the mass difference between reporter ions to generate the same precursor ion. (B) During fragmentation the bond between the reporter ion and balance part is cleaved, permitting relative quantification based on the intensity of the reporter ion. Adapted from [124].	26
1.16	'Light' (top) and 'heavy' arginine (bottom)	27
1.17	Workflow for iTRAQ-based quantification of proteins	28
1.18	Schematic overview of the principle behind LESA	30
1.19	The commercial available LESA system (TriVersa Nanomate) from Advion Biosciences (Ithaca, NY)	31

1.20	Chip used for nanoESI in LESA-MS/MS analysis . . . . .	31
1.21	Schematic representations of ionisation theories for ESI. (A) The <i>ion evaporation model</i> . The droplet evaporates until the charge density reaches the Rayleigh limit (B) The <i>charge residue model</i> . Every droplet contains a single analyte from which the ion is released once the solvent has evaporated. Adapted from [169] . . . . .	33
2.1	(A) Schematic representation of the dimensions of a spot on the Droplet Microarray. Superhydrophilic spot size: 2.8 mm; superhydrophobic border width: 1.7 mm (B) Macroscopic image of a 10 mM hemin solution dispensed ( $V = 3 \mu\text{L}$ ) onto glass (top) and DMA (bottom). Images reveals the trapping of the aqueous solution within the superhydrophobic areas whilst the solution spreads out widely on a microscope glass slide . . . . .	39
2.2	Signal intensity before Savitzky-Golay smoothing (black), after Savitzky-Golay smoothing (red) and after signal correction; Left: base peak (creatinine; $m/z$ 114.0663); Right: peak at $m/z$ 203.0529 ( $\text{C}_6\text{H}_{12}\text{O} [\text{M}+\text{Na}]^+$ ). It can be seen that after correction the signal intensity is restored at its original level . . . . .	42
2.3	Depositing and extraction of samples from the Droplet Microarray and glass surface. (A) 10 mM hemin in 50 mM NaOH manually pipetted onto glass (top) and Droplet Microarray (bottom). Sample volume: $3 \mu\text{L}$ ; spot size DMA: $\sim 2.8 \text{ mm}$ ; spot size glass: $\sim 4.0 \text{ mm}$ . (B) Microscope image (Optical Profile, Zeta Instruments; $5\times$ magnification) of dried urine on Droplet Microarray (left) and glass (right). Images taken before (top) and after extraction (bottom). Visually higher efficient extraction from DMA as can be derived from the empty spot in the microscope image. Sample volume: $2 \mu\text{L}$ . . . . .	44

2.4	PCA scores plot for included ions found in urine before (green circles) and after tea intervention (orange squares). (A) LESA-MS on glass surface ( $R^2 = 0.7184$ ; $Q^2 = -0.1551$ ). (B) LESA-MS on DMA ( $R^2 = 0.9609$ ; $Q^2 = 0.8938$ ) . . . . .	46
2.5	Volcano plots for common ions between control and tea intervention urine samples. Green: significantly increasing ion after tea intervention. Red: significantly decreasing ion after tea intervention. (A) Analysis from glass substrate. (B) Analysis from DMA . . . . .	47
3.1	Visualisation of the rolling droplet technique as described by Popova <i>et al.</i> [188]. The extreme difference in wettability of the pattern allows formation of individual droplets on the array . . . . .	55
3.2	Optimising solvent volume for LESA-MS/MS analysis of tryptic peptides. (A) TIC deviation expressed as CV. (B) Average TIC per solvent volume . . . . .	56
3.3	The impact of extraction time on the intensity of the tryptic BSA peptide QTALVELLK (2+; $m/z$ 507.8138) . . . . .	58
3.4	Comparison of the signal stability for different extraction solvents. (A) Jitter plot for the average TIC per extraction solvent. (B) Statistical analysis (ANOVA + Tukey-Kramer <i>post hoc</i> analysis) for comparison of the TIC intensity. (C) TIC variation for different extraction solvents. (D) Statistical analysis (ANOVA + Tukey-Kramer <i>post hoc</i> analysis) for comparison of the TIC variation. . . . .	59

3.5	Comparison of BSA peptide signal intensity for different extraction solvents. (A) Jitter plot for peptide signal intensity distribution. (B) ANOVA followed Tukey-Kramer <i>post hoc</i> analysis for comparison of peptide signals. Non-overlapping intervals are significantly different. (C) Jitter plot for variation in peptide signal intensity. (D) ANOVA followed Tukey-Kramer <i>post hoc</i> analysis for comparison of variation in peptide signal. Non-overlapping intervals are significantly different . . . . .	60
3.6	Jitter plot for comparison of the yeast peptide intensities for FA and TFA as extraction solvent additive . . . . .	62
3.7	Investigating the benefit of DMSO as additive for the LESA extraction solvent. (A) Peptide intensity distribution with and without DMSO. No difference in signal intensity was observed with or without 1% (v/v) DMSO (Student's <i>t</i> -test : $p = 0.0827$ ). (B) Peptide intensity variation with and without DMSO. Signal repeatability significantly improved using 1% (v/v) DMSO as an additive (Levene's test: $p < 0.001$ ) . . . . .	63
3.8	The LESAProteomics MATLAB class for processing of proteomics data acquired via LESA-MS/MS . . . . .	65
3.9	Annotated MS <sup>1</sup> spectrum for BSA peptides. QTALVELLK is indicated by the black arrow . . . . .	66
3.10	Annotated MS <sup>2</sup> spectrum for the most intense BSA peptide (QTALVELLK 2+) . . . . .	67
3.11	Jitter plot for comparing (A) MS <sup>1</sup> and (B) MS <sup>2</sup> quantification . . . . .	67
3.12	Annotated MS <sup>1</sup> spectrum for tryptic peptides generated from recombinant human insulin. The most intense peak ( $m/z$ 430.2212) represents GFFYTPK (2+) whilst the other annotation is the singly charged peptide ion for GFFYTPK (1+; $m/z$ 859.4342) . . . . .	69

3.13	Comparing <i>in situ</i> digestion performance for recombinant human insulin with different proteases. (A) Jitter plot for the peptide intensity distribution of GFFYTPK. (B) ANOVA with subsequent <i>post hoc</i> analysis (Tukey-Kramer test) for statistical comparison of peptide intensities. Non-overlapping intervals are significantly different . . . . .	70
3.14	The addition of DMSO prevented the evaporation of the digestion buffer (B) compared to the standard 100 mM AmBic digestion buffer. After 2 hours incubation at 37°C , the deposited 2 µL is still present in B (indicated by the red circles) whilst in A no solution is visible. This indicates the addition of DMSO is required to prevent the digestion to stop before completion Scale bar: 5 mm . . . . .	71
3.15	Investigating the effect of DMSO as an additive to prevent evaporation of the digestion buffer. (A) Jitter plot for comparison of the GFFYTPK peptide intensity. (B) Jitter plot for comparison of the digestion efficiency measured by the peptide/intact protein ratio . . . . .	71
3.16	Comparison of the digestion efficiency for insulin under different digestion conditions. (A) Intensity (peak height) distribution for GFFYTPK 2+ for each condition. (B) ANOVA followed by Tukey's <i>post hoc</i> correction for statistical comparison of the peptide intensities. Non-overlapping intervals are significantly different. (C) Distribution of the digestion efficiency for insulin (peptide-intact protein ration) for each digestion condition. (D) ANOVA followed by Tukey's <i>post hoc</i> correction for statistical comparison of the digestion efficiency. Non-overlapping intervals are significantly different. . . . .	73
3.17	Annotated MS <sup>1</sup> spectrum for tryptic peptides generated from transferrin. The arrow indicates the most intense peptide (DSGFQMNQLR), which is shown in detail in the figure insert . . . . .	76



3.18	Fragmentation spectra used for MS <sup>2</sup> quantification. (A) Most intense peptide for insulin: GFFYTPK. (B) Most intense peptide for transferrin: DSGFQMNQLR . . . . .	77
4.1	Schematic representation of the humidity chamber used for <i>in situ</i> digestion . . . . .	83
4.2	Structures for selected monomers for testing the Droplet Microarray as printing substrate . . . . .	89
4.3	Schematic representation of print layout for testing the Droplet Microarray as printing substrate . . . . .	89
4.4	Representative TOF-SIMS ion images for the selected monomers before seven days of incubation in deionised H <sub>2</sub> O . (A) tBCHA (B) iBMA (C) DEGDMA (D) NGPDA. Chemical moieties detected by TOF-SIMS are highlighted in orange on the chemical structures . . . . .	91
4.5	Representative TOF-SIMS ion images for the selected monomers after seven days of incubation in deionised H <sub>2</sub> O . (A) tBCHA (B) iBMA (C) DEGDMA (D) NGPDA. Chemical moieties detected by TOF-SIMS are highlighted in orange on the chemical structures . . . . .	93
4.6	Structures used for second optimisation experiment . . . . .	95
4.7	Chemical structures for unsuccessfully printed monomers. (A) glycidyl acrylate (GA), (B) butyl acrylate (BA), (C) ethylene glycol diacrylate (EGDA), (D) ethylene glycol dicyclopentyl ether acrylate (EGDPEA), (E) ethyl acrylate (EA), and (F) acrylamide (AAm) . . . . .	97
4.8	Testing different solvent systems for printing monomer solutions onto the Droplet Microarray. (A) Photograph of printed array directly after completion of the printing program. (B) TOF-SIMS ion images for selected polymers for test printing. Representative ions were selected from an in-house TOF-SIMS library . . . . .	99

4.9	Fragmentation spectra for selected peptide GFFYTPK for for identification of insulin . . . . .	103
4.10	Fragmentation spectra for selected peptide DSGFQMNQLR for identification of transferrin. . . . .	104
4.11	Fragmentation spectra for selected peptide YTSWYVALKR for identification of FGF-2 . . . . .	104
4.12	Fragmentation spectra for selected peptide WIHEPK for identification of TGF- $\beta$ 1 . . . . .	105
4.13	Average cell count versus pluripotency maintenance after microarray screening. The use of different polymer chemistry resulted in a change of cell attachment and pluripotency maintenance. Data from Nasir <i>et al.</i> [59] . . . . .	106
4.14	Sequence for printing selection monomer onto the Droplet Microarray ( $\varnothing$ 1.414 mm) . . . . .	107
4.15	Effect of data pre-processing steps on variability between replicates. (A) Pair-wise comparison of all log-transformed data for insulin adsorption. (B) Pair-wise comparison after quantile normalisation and outlier filtering using Thompson's tau test. Improvement in agreement between replicates is observed, i.e. data clean-up led to improved correlation between replicate (better agreement) . . . . .	108
4.16	Non-linear regression results for all cell data and insulin adsorption. The high amount of zero values is masking the true trend in the data whilst showing high correlation . . . . .	109
4.17	Scatter plots for cell attachment as a function of Essential 8 <sup>TM</sup> protein adsorption. (A) Insulin, (B) transferrin, (C) FGF-2, and (D) TGF- $\beta$ 1 . All correlation coefficients were $< 0.001$ and none of the regression models was significant ( $p > 0.05$ ) . . . . .	110

4.18	Prediction of Essential 8 <sup>TM</sup> protein adsorption using ToF-SIMS ion intensities and PLS regression. (A) insulin (B) transferrin (C) FGF-2 and (D) TGF- $\beta$ 1 . The black line represents the identity ( $x = y$ ) . . . . .	112
4.19	Chemical structure of ROCK inhibitor Y-27632 . . . . .	115
5.1	Chemical structures of polymers. (A) BA, (B) THFuA, (C) TCDMDA . . .	118
5.2	Schematic workflow for performing quantitative LESA-MS/MS analysis for proteins adsorbed onto polymer-coated 6-well plates. Cells are seeded and cultured on the selection of polymers and subsequently stained with DAPI and OCT4 to quantify the total number of cells and number of pluripotent cells respectively. Cell quantification was performed using fluorescence microscopy and Cell Profiler [61] for automated quantification. In parallel, protein adsorption experiments were done on the same selection of polymers. A spectral library is first generated for the Essential 8 <sup>TM</sup> proteins. The identified peptide ions were used to create a list of target peptides allowing $\sim 1$ minute analysis per sample. MS/MS spectra were for the sample spectra were compared against the library. Then, protein quantification was done using the sum of the identified peptide ion intensities. To assess difference in protein adsorption between different polymer, ANOVA was used. . . . .	121
5.3	Comparison of number of identified peptides between ambient <i>in situ</i> digestion and microwave-assisted <i>in situ</i> digestion. †: $p < 0.05$ . . . . .	122
5.4	Comparison of sequence coverage between ambient <i>in situ</i> digestion and microwave-assisted <i>in situ</i> digestion . . . . .	123
5.5	Experimental design for quantitative proteomics on a selection of synthetic substrates . . . . .	123
5.6	Identified tryptic peptide through library search for insulin . . . . .	124
5.7	Identified tryptic peptide through library search for transferrin . . . . .	125
5.8	Identified tryptic peptide through library search for TGF- $\beta$ 1 . . . . .	125

5.9	Identified tryptic peptide through library search for FGF-2 . . . . .	126
5.10	Comparison of protein adsorption for polymer-coated 6-well plates incubated with Essential 8 <sup>TM</sup> medium. Data presented as mean values $\pm$ standard deviation . . . . .	127
5.11	Statistical comparison (ANOVA + Tukey's <i>post hoc</i> analysis) of protein adsorption on different synthetic polymer substrates. (A) Transferrin (B) Insulin (C) FGF-2 and (D) TGF- $\beta$ 1 . Non-overlapping intervals are significantly different . . . . .	128
6.1	Sequential mass spectrometry analysis of pediatric ependymoma tissue microarrays. (A) The tumour tissue was removed, the tumour area was marked and then paraffin-embedded for long term storage. (B) For MS analysis, a TMA block from the archive was sectioned and mounted onto a glass substrate followed by a xylene wash to remove the paraffin. (C) Paraffin-free samples were then analysed by 3D OrbiSIMS followed by (D) LESA-MS/MS. (E) Ions were selected from the mass spectra and aligned. (F) All matrices with ion intensities were then combined (low-level data fusion). (G) Subsequently, data were subjected to partial-least squares-discriminant analysis (PLS-DA) to identify discriminative features in tumour recurrence. (H) Molecular formulae were assigned to the significant ions using the Human Metabolome Data-base. (I) Ions with a putative ID were then submitted to MetExplore for metabolic pathway analysis to identify affected pathways and corresponding genes. . . . .	140

6.2	Investigating the effect of 3D OrbiSIMS analysis on subsequent LESA-MS/MS analysis. The signal intensity of features identified in LESA-MS spectra did not seem to be affected by prior 3D OrbiSIMS analysis. For one patient ( $n = 3$ ), the analysis was carried out with and without OrbiSIMS. The scatter plot reveals a good correlation between obtained feature intensities. Furthermore, Student's $t$ -test revealed no significant difference in signal intensity distribution . . . . .	141
6.3	Identifying putative metabolites in 3D OrbiSIMS and LESA-MS spectra. (A) In total, more ions were identified in the LESA-MS spectra, though the number of ions in negative mode identified as metabolite was higher for SIMS. (B) For all identified ions, the class as described in the Human Metabolome Database was obtained to identify which classes can be detected with each technique. The patched areas represent structurally similar classes. (C) A Venn diagram was derived with data from both surface MS techniques, identifying common and unique metabolite classes that could therefore provide complementary and additive information. . . . .	142
6.4	Identification of significant ions from fused LESA-MS/MS and 3D OrbiSIMS data. (A) PLS-DA scores plot reveals clustering of patients based on tumour recurrence. (B) Box plot for significant ions ( $p < 0.05$ ) identified using Student's $t$ -test and FDR estimation through a permutation test. . . . .	144
6.5	Ion images for L-phenylalanine ( $m/z\ 146.0610 \pm 0.001\ m/z$ ) for patients in the <i>eventual relapse</i> group . . . . .	148
6.6	Ion images for L-phenylalanine ( $m/z\ 146.0610 \pm 0.001\ m/z$ ) for patients in the <i>no relapse</i> group . . . . .	148

6.7	Comparison of intra-tumour heterogeneity between the <i>no relapse</i> and <i>eventual relapse</i> subgroups. Representative 3D OrbiSIMS images for adenine in the (A) <i>eventual relapse</i> and (B) <i>no relapse</i> group. The dynamic range for adenine was significantly smaller compared to L-phenylalanine (ANOVA: $p = 0.0039$ ). . . . .	149
6.8	Investigating of difference in intra-tumour heterogeneity between <i>no relapse</i> and <i>eventual relapse</i> subgroups. No difference in heterogeneity (dynamic range) was observed. . . . .	150
6.9	Statistical comparison for the spatial abundance of L-phenylalanine between ependymoma subgroups. (A) The number of pixels (ion frequency) in which L-phenylalanine was detected was higher for the <i>no relapse</i> group (Mann-Whitney: $p = 5.29 \times 10^{-4}$ ). (B) In addition, the average pixel intensity for L-phenylalanine was found to be higher in the <i>no relapse</i> group (Student's <i>t</i> -test : $p = 2.12 \times 10^{-5}$ ) . . . . .	151
6.10	Gene expression analysis of primary ( $n = 72$ ) and recurrent ependymoma ( $n = 47$ ). The dendrogram reveals two distinct groups for which most of the recurrent samples belong to the first group and most of the primary samples to the second group. The samples associated with strong evidence for the association between the response (primary vs. recurrent) and the gene expression profile of the gene set (10 genes) have small p-values (tall bars in the bottom plot) . . . . .	153

# List of Tables

1.1	Identified TOF-SIMS ions affecting colony formation [32] . . . . .	8
1.2	Overview of commercially available proteases for MS-based proteomics	18
1.3	Overview of commonly used LC types for proteomics . . . . .	19
1.4	Available search algorithms for peptide identification from MS/MS spectra	24
2.1	Components of the standard mixture . . . . .	39
2.2	Reproducibility of signal intensities for every standard mixture component extracted from glass and Droplet Microarray . . . . .	43
2.3	Statistical comparison of LESA extraction reproducibility from glass and DMA . . . . .	45
3.1	Solvent volume parameters used for optimisation . . . . .	51
3.2	Signal intensity for hydrophobic BSA peptides when analysed with 100 mM NH <sub>4</sub> OH compared to ACN/H <sub>2</sub> O 1:1 (v/v) . . . . .	61
3.3	Performance of search algorithms for identifying peptide from a yeast digest analysed by LESA-MS/MS . . . . .	64
3.4	Comparison of MS protein quantification strategies . . . . .	67
3.5	Identification rate of target proteins using DDA and DIA . . . . .	68
3.6	Digestion conditions for used proteases . . . . .	70
3.7	Digestion conditions used for the second round of experiments using DMSO/100 mM AmBic 1:9 v/v . . . . .	72
3.8	Variation in digestion efficiency per protein for different digestion conditions . . . . .	74

3.9	Comparison of quantification strategies for insulin and transferrin . . . .	77
4.1	Amino acid sequences for target proteins in FASTA format . . . . .	85
4.2	Comparison of TOF-SIMS ion intensities before and after wash . . . . .	94
4.3	Overview of detected and not detected polymers using TOF-SIMS . . . .	96
4.4	Identified PSMs per Essential 8 <sup>TM</sup> protein . . . . .	100
4.5	Number of outliers (Thompson tau) and removed polymers (abundance filter) per Essential 8 <sup>TM</sup> protein after filtering . . . . .	107
4.6	Comparison of protein adsorption values for Essential 8 <sup>TM</sup> proteins across different polymers . . . . .	108
4.7	Results of pre-processing protein adsorption data for building a predict- ive PLS model . . . . .	111
4.8	Results of permutation test for predictive PLS models . . . . .	113
4.9	Important ions in PLS model for predicting Essential 8 <sup>TM</sup> protein ad- sorption. Reported values are <i>m/z</i> values . . . . .	113
6.1	Demographic information for patients included in the analysis ( <i>N</i> = 7; <i>n</i> = 3) . . . . .	135
6.2	Annotations for significant ions ( <i>p</i> < 0.05). Fold changes were calculated by dividing the average ion intensity of the <i>no relapse</i> group by the aver- age ion intensity in the <i>eventual relapse</i> group . . . . .	144
6.3	Putative metabolite IDs for significantly affected pathways ( <i>p</i> < 0.05) between ependymoma sub groups. Pathway analysis was performed using MetExplore . . . . .	146



# Abstract

**S**URFACE-based mass spectrometry analysis benefits from the minimum sample preparation required and high throughput nature of the analysis (few minutes per sample) and shows therefore potential for tackling current issues in the field of proteomics and metabolomics. This thesis aims to develop a robust high throughput methodology for the quantitative analysis of surface-adsorbed proteins and for untargeted metabolomics

One of the current problems in the field of biomaterials research is the limited understanding of the mechanistic behind cell attachment and behaviour on polymeric substrates. Fully synthetic substrates have been identified which support growth and survival of human pluripotent stem cells. Pluripotent stem cells are a valuable cell type for regenerative medicine due to their ability to differentiate into the three germ layers. To treat a single patient, more than a billion stem cells are required. Current cell systems use biological feeder layers for stem cell expansion. However, these animal-derived matrices are expensive, undefined, and show high batch-to-batch variation. In order to move towards reproducible, industrial culturing of stem cells a suitable growth substrate needs to be defined. Through high throughput biomaterials discovery, it was shown that some fully synthetic polymers can maintain stem cell cultures to a similar level as biological substrates.

Current understanding of the response of cells on those synthetic polymers is relatively poor. Research has shown that coating of synthetic polymers with culture medium-derived proteins increase the cell attachment which is required for cell survival. This shows the potential role of culture medium proteins in the response mech-

anism of cells on synthetic polymers. However, current technology does not allow analysis of (combinatorial) polymer libraries which has limited the understanding of relation between cell response and physicochemical properties and molecular features of the polymers. A full understanding of the cell-polymer response mechanism would allow the development and rationalisation of synthetic polymers for culturing of pluripotent stem cells.

It was shown that liquid extraction surface analysis-tandem mass spectrometry (LESA-MS/MS) is a suitable analytical technique for the analysis of *in situ* digested proteins. LESA is a commercial system which can automatically extract analytes from a given substrate and directly introduce the sample into the MS. Here, this potential was further explored for polymer array screening as well as polymers taken forward for scale-up experiments. A suitable substrate was chosen (Droplet Microarray) which allowed control over the spreading of the monomer solutions, digestion solution, and organic extraction solvent for reproducible MS results. With carefully optimised LESA and MS parameters, difference in protein adsorption could be detected between different chemical surfaces. These difference in protein adsorption did not show a good correlation with the observed cell response (attachment and number of pluripotent stem cells). Through multivariate modelling was found that surface chemistry was found to play a role in protein adsorption. Whilst array screening did not reveal solid evidence of the importance of protein adsorption in relation to cellular response, experiments of protein adsorption on a larger surface area (6-well plates) revealed higher protein adsorption on polymers with higher numbers of pluripotent stem cells. Altogether, LESA-MS/MS shows to be an interesting tool to quantitatively assess protein adsorption on synthetic polymers. The developed methodology can not only be further used to study more complex growth media for human cell lines, but also extended study the relation between protein adsorption and response of different organisms. The addition of LESA-MS/MS to high throughput screening of material microarrays might reveal vital information and could assist in proper choice of polymers for biomedical purposes.

Further interest of surface analysis comes from the field of oncometabolomics. In this thesis, paediatric ependymoma were analysed by Orbitrap secondary ion mass spectrometry (3D OrbiSIMS) and LESA-MS/MS. The main challenge here was to acquire data using only minimal tumour tissue which was available in the form of a tumour tissue microarray. By analysing the same tumour tissue with two complementary mass spectrometry techniques, a more complete metabolite profile could be obtained. Moreover, the combination of 3D OrbiSIMS and LESA-MS/MS data followed by partial-least squares discriminant analysis (PLS-DA) permitted the classification of tumour tissue based on eventual recurrence. This means that certain metabolite levels are indicative of tumour relapse. Understanding these changes in metabolite abundance along with the changes in corresponding metabolic pathways could open new insight into ependymoma relapse. Further, this analytical strategy would be suitable to study other types of (tumour) tissues.

# Papers coming forward from this thesis

Meurs, J.; Alexander, M. R.; Levkin, P. A.; Widmaier, S.; Bunch, J.; Barrett, D. A.; Kim, D.-H. Improved Extraction Repeatability and Spectral Reproducibility for Liquid Extraction Surface Analysis–Mass Spectrometry Using Superhydrophobic–Superhydrophilic Patterning. *Anal. Chem.* **2018**, 90 (10), 6001–6005, doi 10.1021/acs.analchem.8b00973.

Nasir, A.; Thorpe, J.; Burroughs, L.; Meurs, J.; Pijuan-Galito, S.; Irvine, D. J.; Alexander, M. R.; Denning, C. Discovery of a Novel Polymer for Xeno-Free, Long-Term Culture of Human Pluripotent Stem Cell Expansion. *Adv. Healthc. Mater.* **2020**, doi: 10.1002/adhm.202001448.

Meurs, J.; Scurr, D. J.; Lourdasamy, A.; Storer, L. C. D.; Grundy, R. G.; Alexander, M. R.; Rahman, R.; Kim, D.-H. Sequential 3D OrbiSIMS and LESA-MS/MS-Based Metabolomics for Prediction of Brain Tumor Relapse from Sample-Limited Primary Tissue Archives. *Anal. Chem.* **2021**, 93 (18), 6947–6954, doi 10.1021/acs.analchem.0c05087.

Meurs, J.; Nasir, A., Pijuan-Galitó, S.; Burroughs, L.; Abdelrazig, S.; Barrett, D.A.; Denning, C.D.; Kim, D.-H.; Alexander, M.A.. Biointerphase and Polymer Chemistry Instruct Pluripotency Maintenance in Human Induced Pluripotent Stem Cells. *Adv. Mater.* (in preparation)

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Biomaterials . . . . .	2
1.2	High throughput biomaterials discovery . . . . .	3
1.3	High throughput screening assays . . . . .	6
1.4	Analysing biomaterial surfaces for investigating cellular response . . . .	9
1.4.1	Time-of-flight secondary ion mass spectrometry (TOF-SIMS) . . .	9
1.4.2	Raman spectroscopy . . . . .	10
1.4.3	X-ray Photoelectron Spectroscopy (XPS) . . . . .	11
1.4.4	Water contact angle . . . . .	12
1.4.5	Data analysis . . . . .	13
1.5	Understanding cell-instructive properties for improved biomaterial design	14
1.5.1	Quantification of cell response . . . . .	14
1.5.2	Current understanding of cell-synthetic material interactions . .	15
1.5.3	Limitations in high throughput screening of proteins adsorbed on biomaterial surfaces . . . . .	16
1.6	Mass spectrometry-based proteomics . . . . .	17
1.6.1	The standard MS workflow for identification and quantification .	17
1.6.2	Limitations of standard mass spectrometry-based approaches for high throughput array screening . . . . .	29
1.7	Ambient mass spectrometry for high throughput screening . . . . .	29
1.7.1	The definition of ambient mass spectrometry . . . . .	29

1.7.2	Applications of ambient mass spectrometry in proteomics . . . .	30
1.8	LESA-MS/MS as a tool for high throughput quantitative surface proteo- mics . . . . .	30
1.9	Aims & objectives thesis . . . . .	34
<b>2</b>	<b>Improving LESA-MS sampling reproducibility through a superhydrophilic- superhydrophobic substrate</b>	<b>36</b>
2.1	Introduction . . . . .	36
2.2	Materials & Methods . . . . .	38
2.2.1	Sample preparation . . . . .	38
2.2.2	Liquid extraction surface analysis-mass spectrometry . . . . .	40
2.2.3	Data analysis . . . . .	41
2.3	Results & Discussion . . . . .	42
2.3.1	Robust LESA with Droplet Microarray . . . . .	42
2.4	Improved untargeted metabolomics with Droplet Microarray . . . . .	45
2.5	Conclusion . . . . .	47
<b>3</b>	<b>Optimisation and development of LESA-MS/MS and digestion parameters for <i>in situ</i> quantification of proteins</b>	<b>48</b>
3.1	Introduction . . . . .	48
3.2	Aims objectives . . . . .	50
3.3	Materials & methods . . . . .	50
3.3.1	Reagents chemicals . . . . .	50
3.3.2	Optimising LESA extraction parameters . . . . .	51
3.3.3	Inferring protein identities . . . . .	53
3.3.4	Label-free protein quantification . . . . .	53
3.3.5	MS acquisition strategies . . . . .	54
3.3.6	Reproducible <i>in situ</i> digestion . . . . .	54
3.3.7	Quantification of target proteins . . . . .	54
3.3.8	Analysis . . . . .	55

3.4	Results & discussion . . . . .	55
3.4.1	Optimising LESA & MS parameters . . . . .	55
3.4.2	Search algorithm selection for maximum peptide identification .	63
3.4.3	Reproducible and accurate protein identification & quantification	64
3.4.4	Robust <i>in situ</i> digestion . . . . .	68
3.4.5	Protein quantification with optimised parameters . . . . .	74
3.5	Conclusion . . . . .	78
<b>4</b>	<b>High throughput screening of acrylate and methacrylate homopolymers for examining adsorption of culture medium proteins in relation to cell response of human induced pluripotent stem cells (hIPSCs)</b>	<b>79</b>
4.1	Introduction . . . . .	79
4.2	Materials & Methods . . . . .	81
4.3	Reagents & chemicals . . . . .	81
4.3.1	Array printing . . . . .	81
4.3.2	TOF-SIMS . . . . .	82
4.3.3	Incubating microarrays in Essential 8™ . . . . .	82
4.3.4	<i>In situ</i> digestion . . . . .	82
4.3.5	LESA-MS/MS for spectral library generation . . . . .	83
4.3.6	LESA-MS/MS for high throughput protein quantification on poly- mer microarrays . . . . .	84
4.3.7	Protein identification & quantification . . . . .	84
4.3.8	Microarray screening of human induced pluripotent stem cells .	86
4.3.9	Data cleaning & statistical analysis . . . . .	86
4.4	Results & Discussion . . . . .	88
4.4.1	Optimisation of printing parameters . . . . .	88
4.4.2	Building a spectral library for protein identification and quanti- fication . . . . .	100

4.4.3	Essential 8 <sup>TM</sup> protein does not correlate with attachment of pluripotent stem cells . . . . .	105
4.4.4	Investigation of structural related protein adsorption using molecular descriptors and PLS regression . . . . .	111
4.4.5	Potential factors affecting cell response . . . . .	114
4.5	Conclusion . . . . .	115
<b>5</b>	<b>Quantitative screening of protein adsorption on scaled-up polymers using LESA-MS/MS analysis</b>	<b>116</b>
5.1	Introduction . . . . .	116
5.2	Aims & objectives . . . . .	117
5.3	Materials & methods . . . . .	117
5.3.1	Reagents & chemicals . . . . .	117
5.3.2	6-well plate fabrication . . . . .	118
5.3.3	Incubation of polymer-coated well plates . . . . .	118
5.3.4	Protein digestion . . . . .	119
5.3.5	LESA-MS/MS . . . . .	119
5.3.6	Protein identification & quantification . . . . .	120
5.3.7	Statistical analysis . . . . .	121
5.4	Results & discussion . . . . .	122
5.4.1	Comparison of ambient <i>in situ</i> digestion and microwave-assisted <i>in situ</i> digestion . . . . .	122
5.4.2	Identification of Essential 8 <sup>TM</sup> proteins from polymer-coated well plates . . . . .	123
5.5	Conclusion . . . . .	129
<b>6</b>	<b>Sequential 3D OrbiSIMS and LESA-MS/MS-based metabolomics for prediction of brain tumor relapse from sample-limited primary tissue archives</b>	<b>131</b>
6.1	Introduction . . . . .	131
6.2	Aims & Objectives . . . . .	134



6.3	Materials & Methods . . . . .	134
6.3.1	Reagents & chemicals . . . . .	134
6.3.2	Experimental design . . . . .	134
6.3.3	Sample preparation for MS analysis . . . . .	135
6.3.4	3D OrbiSIMS analysis . . . . .	135
6.3.5	LESA-MS/MS analysis . . . . .	136
6.3.6	Ion selection for multivariate analysis . . . . .	137
6.3.7	Multivariate analysis and pathway analysis . . . . .	137
6.3.8	Chemical heterogeneity in 3D OrbiSIMS images . . . . .	138
6.3.9	Statistical analysis . . . . .	138
6.3.10	Gene expression analysis . . . . .	138
6.4	Results & Discussion . . . . .	139
6.4.1	Serial 3D OrbiSIMS and LESA-MS/MS for in situ metabolite pro- filing . . . . .	139
6.4.2	Complementary metabolite profiling with 3D OrbiSIMS and LESA- MS/MS . . . . .	141
6.4.3	Predicting tumor recurrence from metabolite profiles . . . . .	142
6.4.4	Investigating metabolic intra-tumour heterogeneity using 3D Or- biSIMS imaging . . . . .	147
6.4.5	Validation of MS-based metabolomics with publicly available gene expression data . . . . .	151
6.5	Conclusion . . . . .	153
<b>7</b>	<b>Overall conclusion &amp; future perspectives</b>	<b>155</b>
	<b>Appendices</b>	<b>205</b>
<b>A</b>	<b>Acronyms and structural formulas for used monomers</b>	<b>205</b>

# Introduction

## 1.1 BIOMATERIALS

**B** IOMATERIALS play a main role in nowadays medicine. A biomaterial is defined as a non-drug substance or combination of substances derived from natural or synthetic sources which can serve as treatment, augmentation or substitute for any malfunction of the body. Application examples of biomaterials are: drug delivery, support for tissue regeneration and surface coatings for e.g. infection resistance in medical devices or support of implants [1–4]. Biomaterials can e.g. serve as a (temporary) replacement for a malfunctioning body part [5–7] as well as system for increasing efficacy of treatments [8–10].

Finding new biomaterials is essential to meet nowadays clinical questions as antibiotic resistance [3, 11]. It is therefore important to investigate in a fast and efficient manner which materials are potential candidates to be used. Most biomaterials research has been using a trial-and-error approach to identify candidate materials limiting the discovery of new materials [12]. With current available techniques and computational power, data-driven high throughput discoveries are emerging[12–15] .

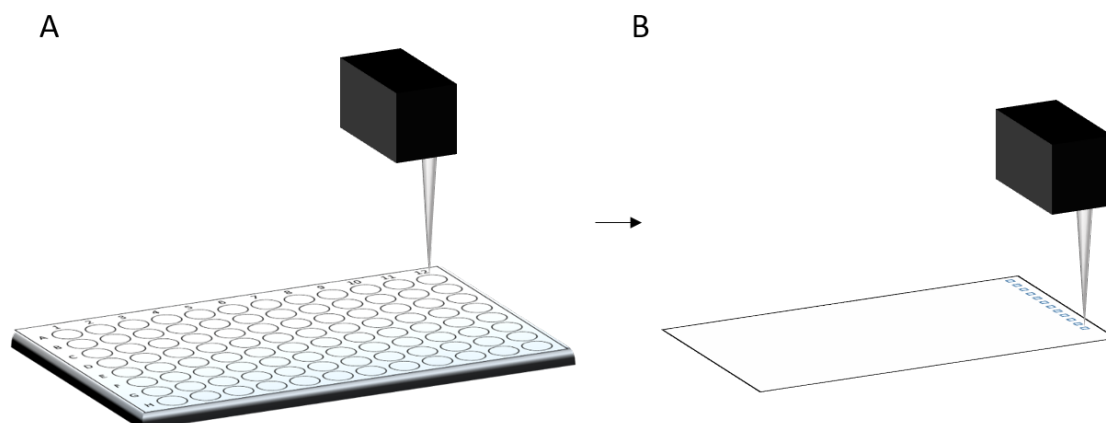
An important part in the discovery of new biomaterials is investigating the interaction between the material of interest and endogenous and exogenous compounds. Throughout the years, development of new biomaterials has been limited due to the low throughput nature of assays for studying interactions between cells and biomater-

ials [1]. The idea behind high throughput biomaterials discovery is to use miniaturised libraries to identify relations between the chemical structure and cellular response [13]. Understanding how these interactions work will be a major leap towards rationalising and improving design of new fit-for-purpose biomaterials. High throughput strategies have already led to the development of an anti-bacterial coating for urinary catheters [16]. In order for high throughput biomaterials discovery to be successful, the library of candidate materials should consist of physicochemical diverse materials to maximise the discovery space [17].

### 1.2 HIGH THROUGHPUT BIOMATERIALS DISCOVERY

The surface chemistry of a material plays an important role in the ability of cells to adhere to the surface [18]. Synthetic materials like polystyrene are often used for cell culturing. One of the drawbacks of using these relatively simple polymers is the limited cell adherence and therefore the materials require a protein coating in order to enhance cellular adherence [18]. However, the use of protein coatings can introduce interference to the sample, which can complicate any biological assay. In the quest to discover materials that offer cell adherence and pluripotency maintenance, it is essential that no false discoveries are made due to matrix interference. Proper positive and negative controls are therefore a necessity to include during an assay.

One of the main tools in high throughput biomaterials discovery is the use of microarrays [15, 19]. These microarrays consist of numerous unique polymers which are printed on unique locations onto a substrate [15, 19]. The production of microarrays is mostly accomplished by printing and to some lesser extent by lithography [15] (Figure 1.1).



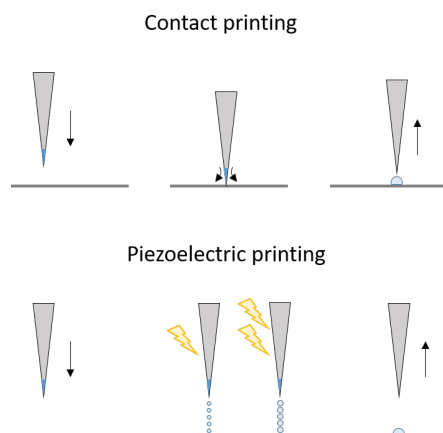
**Figure 1.1:** Schematic representation of a material microarray fabrication. (A) The material solution is collected from the source plate. (B) The solution is then dispensed onto the printing substrate. Adapted from Algahtani *et al.* [15]

Substrate selection is an essential part in microarray production. In general, microarrays are produced by spotting liquid material on a solid material. Suitable substrate material should therefore allow the formation regular-shaped droplets in order to provide high throughput read outs of biological assays [15]. Furthermore, the substrate should ideally not introduce any defect in the materials of interest, establish sufficient adhesion of the material and should be compatible (low fouling [20]) with the analytical method of interest for screening the microarray [15, 19, 20]. Poly(hydroxyethyl methacrylate) (pHEMA) is commonly used as substrate for microarray production since it supports entrapment of the pre-polymer / polymer when introduced on the surface. In addition, pHEMA has shown to reduce the cellular growth rate [15]. An alternative to pHEMA is the use of an agarose coating since it immobilises the pre-polymer/polymer in a similar way as pHEMA and possesses antifouling properties which limits the adherence of cells [15, 21].

As mentioned before, the production of microarrays is in general performed using printing methods [15]. These methods can be divided into contact and non-contact printing (Figure 1.2). In direct contact printing, a solid or grooved pin is loaded with the pre-polymer/polymer (ink). A robot arm then moves to a predefined location and a set volume of ink is released onto the surface once in contact. The main advantages

## CHAPTER 1

of direct contact printing are high reproducibility of spotting, flexibility in choice of solvent, low maintenance and easy to clean [15]. Nonetheless, the pin can easily damage the surface when it is making contact. In addition, the amount ink deposited on the surface is rather difficult to set since one cannot tune the dispensed volume [15]. Both issues can be overcome by using non-contact printing. Non-contact printing can be divided into two main methods, namely piezoelectric printing and thermal printing [15]. In the first method of printing a droplet is released from the nozzle by applying a voltage to the nozzle which is made from a piezoelectric material [15]. By altering the voltage and frequency settings the droplet speed and size can be controlled [15, 22]. The latter non-contact printing method releases solvent from the nozzle by heating of an air bubble causing an increase in pressure inside the solvent chamber which results in ejection of a droplet [15, 23]. Both methods allow deposition of picoliter volumes onto the substrate [15]. Non-contact printing allows on-surface mixing of solvents (e.g. sample and substrate) which can be time-saving compared to pre-mixing of solvents [15]. Further, the non-contact nature of printing avoids contaminating the nozzle with particulates on the surface and avoids damaging the substrate [15]. In addition, non-contact printing methods are able to rapidly print many materials with a spot size which ranges in the low micrometre range [24]. On the other hand, only non-volatile solvents can be used for non-contact printing since volatile solutions would evaporate before reaching the substrate. Also, solvent viscosity is an important parameter since more viscous solvents are prone to block the printing nozzle [15, 25].



**Figure 1.2:** Schematic representation of different printing methods. In contact printing, the solution is released upon contact with the substrate. For piezoelectric printing, a voltage is applied on the piezoelectric pin which causes the pressure to increase within the solvent chamber permitting release (jetting) of picoliter droplet from the pin. The droplet volume can be tuned by altering the applied voltage.

### 1.3 HIGH THROUGHPUT SCREENING ASSAYS

The microarray platform has been used several types of screening. For instance, this approach has been used for drug discovery, gene function screening, cell adhesion assays and material discovery for controlling stem cell fate [26]. Below, only results and findings of cell adhesion and material discovery papers will be revisited, since this is related to the project.

Anderson *et al.* [27] used a range of polyester materials on a pHEMA substrate to investigate how these materials affect stem cells. They observed inhibition of stem cell attachment to a copolymer of poly(lactide), poly(glycolide) and poly(ethylene glycol) (53:21:26 v/v/v) inhibits cell attachment. This is probably due to the presence of poly(ethylene glycol) since this polymer was already found to inhibit cell attachment [27, 28]. The stem cell attachment could be increased by adding a copolymer of poly(lactide) and poly(glycolide) (70:30 v/v). Same experiments were performed to investigate cell attachment of bovine chondrocytes. The average cell count per polymer spot was found to be 120.

A microarray containing extracellular matrix (ECM) proteins has been employed

for investigating cellular differentiation [29]. The proteins used were fibronectin, laminin, collagen I, collagen III and collagen IV. It was found that adherence of hepatocytes was dramatically higher in ECM protein regions compared to regions that did not contain any proteins. Further, the effect of the ECM composition on hepatocyte function was studied by staining for intracellular albumin. Most intense fluorescence was observed when the matrix contains collagen IV. However, additional experiments using only collagen IV as matrix in a serial dilution did not show a significant effect on the function of hepatocytes. Therefore, it was concluded that collagen IV alone is not responsible for the changes in cell function but rather due to a combination of proteins. Further experiments showed that also fibronectin had a positive effect on albumin fluorescence whilst laminin and collagen III had a negative effect when tested individually. Combinations of proteins could either led to increased or decreased albumin secretion.

Another study by Soen *et al.* [30] explored differentiation of neural cells using a microarray approach. As is also described in the previous paragraph, different ECM proteins were tested for their effect on neural progenitor cells. Again, it was observed that ECM proteins (laminin, fibronectin and vitronectin) support cell adherence while in absence of these components cell adherence is barely observed. Differentiation of cells was investigated by staining with TUJ1 and GFAP. Soen *et al.* [30] observed that neural progenitor cells differentiated preferentially to neurons on a laminin substrate. In general, the composition of the substrate controls the differentiation rate to either neurons or glial cells.

Cell adhesion, proliferation and differentiation was also studied by Nakajima *et al.* [31] by printing an array of proteins. Their research also showed that cellular adhesion, proliferation and differentiation are depended on the composition of the growth substrate. Neurospheres showed good adherence to laminin, fibronectin, ProNectin<sup>TM</sup> F, ProNectin<sup>TM</sup> L and poly(ethyleneimine) (MW = 800 Da). Low adherence of neurospheres was observed on collagen type proteins, gelatins and higher molecular weight poly(ethyleneimine). No possible reason for the difference in adherence to poly(ethyleneimine) of different weights was given.

Mei *et al.* [32] used a microarray approach for combinational screening of bacterial attachment to acrylate polymers. It was found that coating the array with fetal bovine serum (FBS) allows efficient cell attachment across the microarray. Also, they found that the number of acrylate groups influences colonial formation of human embryonal stem cells (hESC; Table 1.1). Furthermore, they noted that materials with a low elastic modulus also have a small water contact angle (WCA). By applying time-of-flight secondary ion mass spectrometry (TOF-SIMS) on the incubated polymers, it was found that certain molecular moieties show difference in supporting colony formation. Also, a poor correlation was found between the surface roughness of the material and colony formation.

**Table 1.1:** Identified TOF-SIMS ions affecting colony formation [32]

Ion	Molecular moiety	Effect on colony formation
$C_3H_8N^+$	Tertiary amine	Negative
$C_2H_6N^+$	Tertiary amine	Negative
$CN^-$	Tertiary amine	Negative
$C_4H_9^+$	Tertiary butyl	Negative
$C_2H_3^+$	Hydrocarbon	Positive
$C_3H_3^+$	Hydrocarbon	Positive
$CHO_2^-$	Esters	Positive
$C_3H_3O^+$	Esters	Positive
$C_2H_3O$	Esters	Positive
$C_6H^-$	Cyclic structures	Positive
$C_4H^-$	Cyclic structures	Positive
$C_2H^-$	Cyclic structures	Positive

In a follow-up study, Celiz *et al.* [33] used a multigeneration approach for discovery of new materials for stem cell growth. A copolymer of N-(4-hydroxyphenyl) methacrylamide (HPhMA) and HEMA was used for a scaled-up experiment. HPhMA as polymer showed great performance regarding cell adhesion whilst incorporation of HEMA



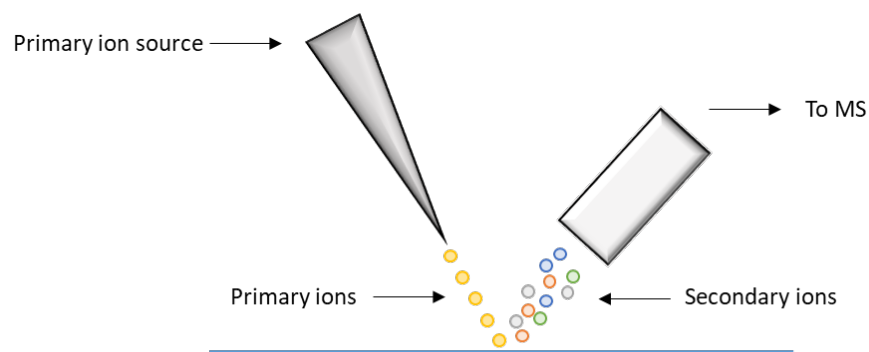
was necessary to stabilise the material. The generated copolymer was stable for at least six months [33].

### **1.4 ANALYSING BIOMATERIAL SURFACES FOR INVESTIGATING CELLULAR RESPONSE**

#### **1.4.1 TIME-OF-FLIGHT SECONDARY ION MASS SPECTROMETRY (TOF-SIMS)**

TOF-SIMS is a surface analysis technique which uses a high energy primary ion beam to sputter components from the surface (Figure 1.3). The primary ion source consists of either a gas cluster ion beam (GCIB) or a liquid metal ion source (LMIS) [34]. Commonly used gas clusters and metals are respectively carbon or argon and bismuth and caesium [35, 36]. Due to high impact energy of the ion source (primary ions; several keV), substrate components are sputtered from the surface and ionised (secondary ions) [34, 37]. Due to the high energy used for the ionisation, secondary ions tend to fragment into smaller ions [38, 39]. During the analysis, the surface is etched by the primary ion beam. Using a low ion dose ( $<10^{12}$  ions $\cdot$ mm $^{-2}$ ), about one monomolecular layer of surface molecules is removed whilst using a higher ion dose results in removal of several nanometers of the surface [39, 40].

TOF-SIMS serves as an excellent tool for surface characterization since libraries of fragments can be created and used for subsequent assignment of unknown substrate composition. Besides the wide usage of SIMS in materials science, there is more and more interest to use SIMS for biological applications such as investigating heterogeneity in cancer cells and studying the metabolome at subcellular level reducing the amount of cells required compared to standard metabolomics analysis with NMR, GC-MS or LC-MS. [41–43].

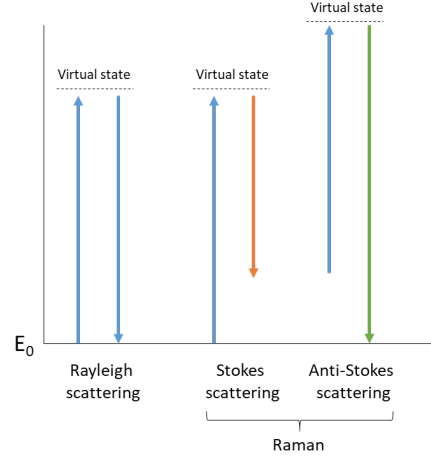


Schematic representation of the principle of SIMS

**Figure 1.3:** Schematic representation of the principle of SIMS

#### 1.4.2 RAMAN SPECTROSCOPY

Raman spectroscopy uses the bond polarisation of molecules to retrieve chemical information. With a laser source, the molecule is excited from the ground state ( $E_0$ ) into a virtual state [44]. The molecule is not fixed in one ground state, but has several different rotational ground states [44]. As shown in Figure 1.4, the excitation as well as the final transition can have respectively different start and end point. In Rayleigh scattering, the molecule returns to the same state after energy transition. Another case of scattering can occur when start state and end state are not the same, so called Stokes and anti-Stokes scattering [44]. To create a spectrum from the energy transitions, the frequency ( $\nu$ ) is converted into a wavenumber. Since the energy transitions are specific for atomic bonds and molecular environment, Raman spectra provide valuable information on the structural composition of a molecule as well as the composition of the analysed matrix [44, 45]. For high throughput screening of biomaterials, Raman spectroscopy offers a complementary analysis strategy to TOF-SIMS for characterisation of the bulk chemistry of the materials on a microarray [15].



**Figure 1.4:** Jablonski diagram electronic states used in Raman spectroscopy. Adapted from [44]

### 1.4.3 X-RAY PHOTOELECTRON SPECTROSCOPY (XPS)

In XPS analysis, the elemental composition of a sample can be determined by measuring the binding energy of photoelectrons [46, 47]. Any element, except for hydrogen and helium, can be detected by this technique [46]. By irradiating a sample with monoenergetic X-rays (photons), electrons in the core shell get excited and, when the energy is large enough, are removed from the core shell (photoelectron). The energy required to remove an electron from the core shell to reach the Fermi level is referred to as the binding energy ( $E_B$ ). XPS measures the loss in kinetic energy ( $E_K$ ) of photoelectron upon removal from which the binding energy can be calculated according to Equation 1.1, in which  $h$  is Planck's constant and  $\nu$  the frequency of the photon, which means that the surface sensitivity is defined by the photoelectron interaction with the matter [48, 49].  $\phi$  represents the work function, which is defined as a instrument-specific correction factor for the energy difference between the Fermi level and the free electron (vacuum) level. A schematic representation of the principle behind XPS is shown in Figure 1.5.

$$E_B = h\nu - (E_K - \phi) \quad (1.1)$$

The binding energy is specific for electronic states of the elements. In addition, since the amount of emitted photoelectrons is proportional to the elemental abundance, XPS can be used for quantitative measurements of the elemental composition of molecules. This makes XPS an excellent tool for characterising materials in terms of their elemental composition.

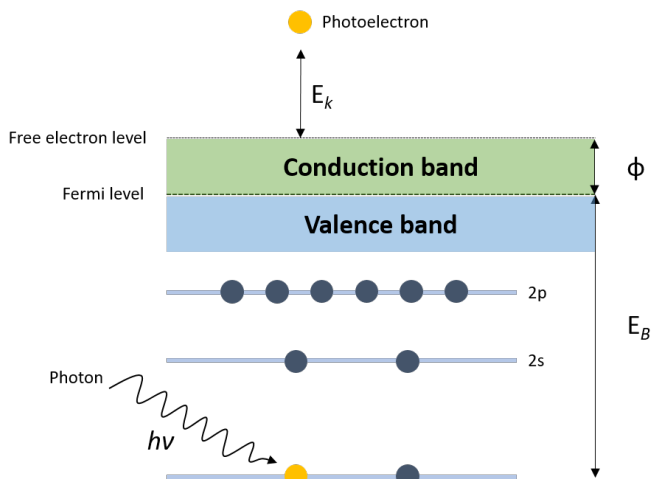
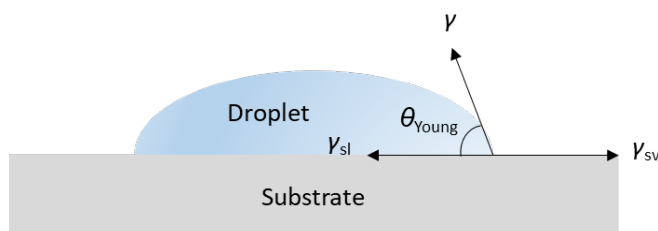


Figure 1.5: Schematic representation of the principle behind XPS. Adapted from [49]

#### 1.4.4 WATER CONTACT ANGLE

WCA measurements provide information on the hydrophobicity of the measured sample. The contact angle is defined as the angle between tangent to the liquid-vapour interface and the sample surface at the three-phase contact line (Figure 1.6) [50]. This surface analysis technique allows measurement on a monolayer ( $\sim 1$  nm) of organic material [51]. For contact angle measurements, the Young equation is used (Equation 1.2) in which  $\gamma_{SV}$  and  $\gamma_{SL}$  are the respective interfacial tensions for the solid-vapor and solid-liquid interface,  $\gamma$  is the surface tension of the liquid and  $\theta_{Young}$  is the Young contact angle [52]. Surface with a higher surface energy show lower contact angles whilst low surface energy results in higher water contact angles [50].

$$\cos \theta_{Young} = \frac{\gamma_{SV} - \gamma_{SL}}{\gamma} \quad (1.2)$$



**Figure 1.6:** Schematic representation of the principle behind water contact angle measurements. Adapted from [50]

#### 1.4.5 DATA ANALYSIS

The data generated by above-mentioned analysis techniques is complex and consists of multiple dimensions. In order to deal with this kind of datasets, sophisticated tools for data analysis are required. These tools will be an aid to visualise and model the data in order to correlate physical and chemical properties of the material with the output from the biological assay. Commonly used multivariate analysis tools are principal component analysis (PCA) and partial least squares (PLS) regression.

PCA is a method for reducing the dimensionality of a dataset [53]. This is done by projecting the data along new axes using eigenvalues representing the largest variance (first principal component) and the second largest variation in the dataset (second principal component) respectively [54, 55]. PCA is an unsupervised data analysis method which means no prior knowledge of the dataset is given to the algorithm and is therefore mainly used for initial exploration of the data [56]. The main use of PCA is to separate data of different classes and find which variables (loadings) are thriving the differences between classes. In terms biomaterials research, PCA can aid to find out which molecular features (loadings) are important in, for instance, protein adsorption to different substrates [57].

In PLS regression, one tries to correlate multiple descriptors (e.g.  $m/z$  values) with their response (e.g. intensity) in order to find a relationship between those two matrices [58]. This multivariate method has often been applied in high throughput biomaterials discovery to, for instance, predict the water contact angle from TOF-SIMS spectra [39] and prediction of cell-material response from TOF-SIMS spectra [32].

## **1.5 UNDERSTANDING CELL-INSTRUCTIVE PROPERTIES FOR IMPROVED BIOMATERIAL DESIGN**

### **1.5.1 QUANTIFICATION OF CELL RESPONSE**

Cell response is defined as the observed behaviour of cells when exposed to a certain condition. For this thesis, the response of interest is the attachment of induced pluripotent stem cells to synthetic polymer surface and the degree of pluripotency maintenance (percentage of number of pluripotent stem cells in the total cell population).

The pluripotent stem cells are commonly cultured in a xeno-free medium (Essential 8™) at 37°C and 5% CO<sub>2</sub> [59]. The cell culture is passaged (harvested and re-cultured) several times to generate a sufficient number of cells. After, the cells are seeded onto the polymer microarrays and incubated for 24 hours at 37°C and 5% CO<sub>2</sub> after which the cell response is investigated [59].

To investigate the cell response of interest, fluorescent stains are available which allow quantification of the cell response through fluorescence microscopy and image analysis in tools as ImageJ [60] or CellProfiler [61]. In order to count the number of cells, the fluorescent stain 4,6-diamidino-2-phenylindole (DAPI) is commonly used [62]. DAPI binds to A-T rich motifs of the DNA present in the cell nucleus. Since cells have only one nucleus, the number of cell nuclei represent the total number of cells.

There are a number of markers available to determine whether a stem cell is still in its pluripotent state. These are Oct4, Nanog and Sox2 which are all transcription factors required to maintain pluripotency [63]. Through immunohistochemical staining the presence of these factors can be determined and, hence, determine whether the cell is pluripotent or not [64]. The same pipeline (fluorescent microscopy and image analysis) can be used to determine the total number of pluripotent stem cells.

### 1.5.2 CURRENT UNDERSTANDING OF CELL-SYNTHETIC MATERIAL INTER-ACTIONS

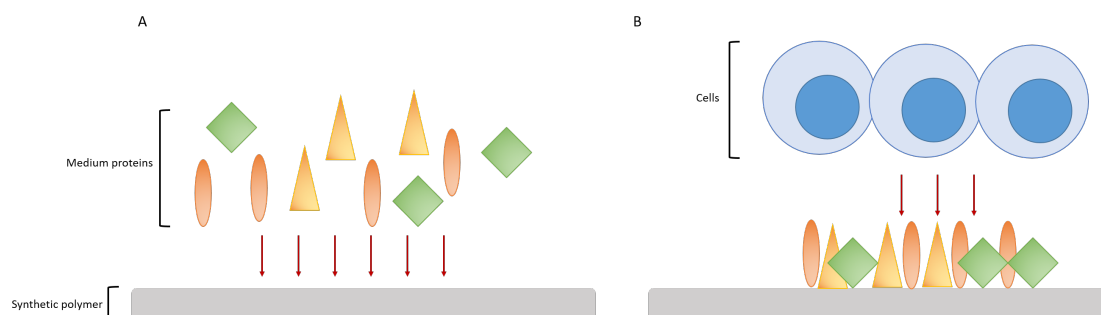
It is known that cells can adhere to biomaterial surfaces [65]. However, it is thought that cells do not directly bind to the material, but to an adsorbed layer of proteins present on the material [65–67]. The degree of protein adsorption and which proteins adsorb are highly dependent on the polymer chemistry and the matrix in which the proteins are present [66, 67]. Hammad *et al.* [67] performed a study to investigate culture medium-derived protein adsorption on polystyrene and plasma-etched polystyrene. Proteins uniquely adsorbing to the plasma-etched surface were identified via liquid chromatography-tandem mass spectrometry (LC-MS/MS). Then those proteins were printed in different concentrations on a *N*-(4hydroxyphenyl) methacrylamide (pHPPhMA) coating, which was found to support human embryonic stem cell (hESC) cultures [33, 67]. When comparing protein-coated polymer to untreated polymer, it was found that certain proteins and protein mixtures increase the cell response (cell attachment) for hESCs. As a result, it can be hypothesised that pre-adsorption of culture medium proteins play an essential role in cell attachment. This in concordance with previous finding that ECM components (proteins and glycosaminoglycans) serve as a anchor site for cells to adhere in order to support cell survival [68, 69]. Furthermore, it is known that cells do not directly adhere to a polymer surface, but require a protein coating in order to adhere [68, 70].

Abdallah *et al.* [69] used LC-MS/MS to identify Matrigel<sup>TM</sup> proteins adsorbing to poly(methyl methacrylate) (pMMA) and pHEMA. Matrigel<sup>TM</sup> is a membrane protein extract derived from Engelbreth-Hom-Swarm mouse tumour cells [71]. They found barely any protein of interest adsorbing to pHEMA, but on pMMA laminin ( $\alpha$ ,  $\beta$ , and  $\gamma$ ) and nidogen 1 were identified as prominent proteins. These proteins were found to promote adhesion and proliferation of epithelial cells [69].

It is further known that fibronectin is an important for cell attachment. Its amino acid sequence has a RGD motif which can serve as cell binding site [13, 72]. By screening an array with a fibronectin RGD peptide gradient, it was found that higher concen-

trations of a RGD peptide led to an increase of smooth muscle cell attachment [13, 73]. In addition, screening gradients of laminin peptides also showed an increase in fibroblast response [13, 74, 75].

From the above-mentioned experimental findings can be derived that culture medium proteins potentially attach first to the synthetic polymer surface and serve as a binding site. Cells can then either attach and proliferate or will lose their viability when attachment is not possible (Figure 1.7). However, due to the limited capability of current techniques to characterise the adherent protein matrix in a high throughput manner, understanding of the mechanistics behind cell-material and protein-material interactions is poor.



**Figure 1.7:** Schematic representation of the potential cell-material response mechanism. (A) First, (a set of) proteins attach to the synthetic polymer surface and serve as (B) a scaffold for the cells to the attach to.

### 1.5.3 LIMITATIONS IN HIGH THROUGHPUT SCREENING OF PROTEINS ADSORBED ON BIOMATERIAL SURFACES

Protein identification and quantification is nowadays performed using MS-based strategies [76]. Studies from Hammad *et al.* [67], Abdallah *et al.* [69], and Tong *et al.* [77] employed LC-MS/MS for characterisation of the adherent matrix on biomaterial surfaces. However, only a small number of surface chemistries (<10) were analysed. Although LC-MS/MS is the golden standard for untargeted proteomics and protein quantification, quantitative screening of protein adsorption on a large library of synthetic polymers is



not compatible due to the required sample preparation and instrumentation time. Furthermore, to quantitatively assess protein adsorption one should understand factors as matrix effects and variation in digestion efficiency due to protein conformation [78, 79]. These factors can have a substantial effect on the accuracy of the quantification. The next section will describe the MS workflow for proteomics in order to understand the current limitations for library screening.

## 1.6 MASS SPECTROMETRY-BASED PROTEOMICS

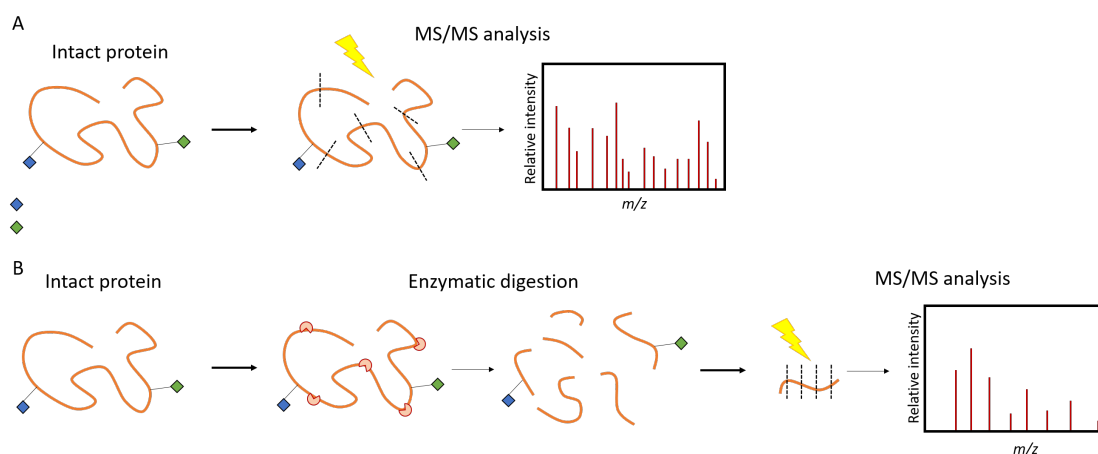
### 1.6.1 THE STANDARD MS WORKFLOW FOR IDENTIFICATION AND QUANTIFICATION

#### 1.6.1.1 TOP-DOWN VS. BOTTOM-UP PROTEOMICS

Though, top-down proteomics has reduced sample preparation steps, in terms of identification and quantification a bottom-up strategy is the preferred method [80]. In top-down mass spectrometry, intact proteins are introduced into the MS and subsequently fragmented (Figure 1.8A). The relatively simple experimental procedure for top-down proteomics allows increased sample throughput and makes it a preferred method for studying intact proteins and protein complexes [80, 81]. However, proteins vary widely in mass ( $\sim 5$ -3000 kDa [82, 83]) which makes top-down analysis of larger proteins within the  $m/z$  range of a mass spectrometer not the preferred strategy of choice for discovery-based proteomics.

For untargeted proteomics, bottom-up proteomics is the preferred method of choice due the higher number of identified proteins identified compared to top-down proteomics [80, 84]. In bottom-up proteomics, proteins are first denatured followed by digestion into peptides using an enzyme (Figure 1.8B). Bottom-up proteomics has shown to be a more sensitive approach than top-down due to improved ionisation efficiency of peptides as well as better ion transmission of peptides compared to intact proteins [85]. In addition, analysis can be performed in a relatively small  $m/z$  range ( $m/z$  400-2000 [86]) since the generated peptides cover a smaller mass range . The most popular

enzyme is trypsin, because of its specific cleavage after arginine and lysine and generation of peptides which fit in a MS-compatible mass range ( $m/z$  400-2000) [86–88]. However, other proteases are available which target other cleavage sites [87]. Proteases and their cleavage sites are shown in Table 1.2.



**Figure 1.8:** Approaches for MS-based protein analysis. (A) Top-down strategy. Intact proteins are introduced into the MS and dissociated via gas-phase collisions. The fragment ions represent a part of the amino acid sequence of the protein. (B) Bottom-up strategy. Intact proteins are first (enzymatically) digested into peptides which are then introduced into the MS and fragmented. The MS/MS spectrum can then be used to identify the peptide based on the amino acid sequence from which a protein identity can be inferred. Adapted from [81]

**Table 1.2:** Overview of commercially available proteases for MS-based proteomics

Enzyme	Cleavage site	Exceptions
Trypsin	C-terminal of K and R	Presence of P
Lys-C	C-terminal of K	-
Lys-N	N-terminal of K	-
rAsp-N	N-terminal of D	-
Thermolysin	N-terminal of L, F, V, A, M and I	-
Pepsin	C-terminal of F, L, Y, and W	-
Elastase	C-terminal of A, V, S, G, L, and I	-
Arg-C	C-terminal of R	-

Information derived from Promega website [89]

## 1.6.1.2 LC-MS/MS

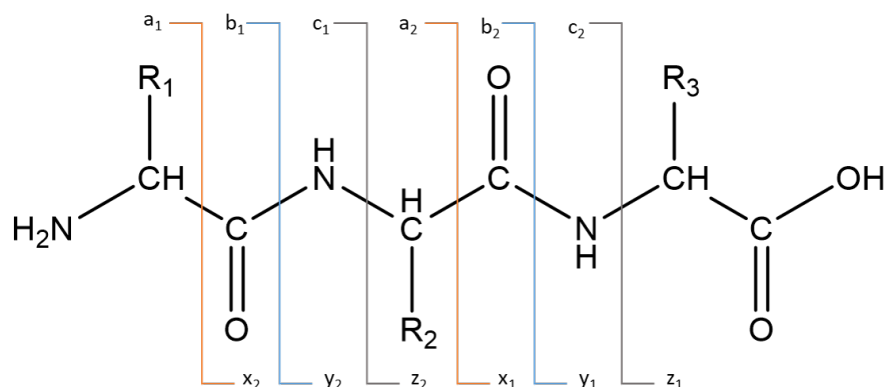
**Chromatography** During a chromatographic gradient, peptides (or other components) are separated based on their affinity to a stationary phase [90]. The stationary phase normally consists of silica-based packing materials (Table 1.3) [91]. For separation of peptides, predominantly reversed-phase (C18) columns are used for their speed and efficiency [91].

**Table 1.3:** Overview of commonly used LC types for proteomics

Chromatography type	Separation mechanism	Proteomics application
C18	Hydrophobicity	Shotgun proteomics
Hydrophilic interaction	Hydrophilic interactions	Hydrophilic peptides
Size exclusion	Size	Peptide purification
Ion exchange	Charge	Shotgun proteomics

Information in table based on [91]

**Fragmentation of peptide ions** Acquisition of tandem MS data of peptides is important for confirming their identity. Fragmentation of ions takes place in a so-called collision cell where the ions are either bombarded with an inert gas or irradiated with a laser. The type of fragmentation used has a severe impact and the ions produced. Roepstorff & Fohlman [92] defined a nomenclature for fragment ions based on the location of bond cleavage caused by the fragmentation which was later updated by Biemann [93]. The nomenclature is explained in Figure 1.9. The *a,b,c*-series always contains the N-terminus of the peptide, whilst the *x,y,z*-series always contains the peptide C-terminus. The numbering starts always from the respective terminus.



**Figure 1.9:** Nomenclature of fragment ions of peptide as defined by Roepstorff & Fohlman and Biemann. Adapted from [94]

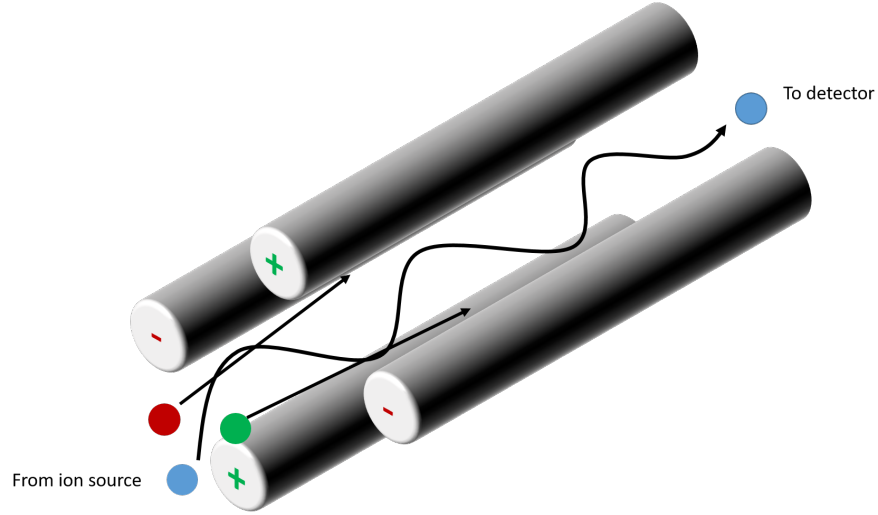
For untargeted (shotgun) proteomics experiments, collision-activated dissociation/collision-induced dissociation/higher-energy collisional dissociation (CAD/CID/HCD) are commonly used. With this type of fragmentation, *b* and *y*-ions are predominately generated, indicating that cleavage takes places at the peptide bond.

**Mass analysers** The mass analyser in the MS is the method of separating the ions by their  $m/z$  value. In proteomics research, five mass analysers are commonly used: TOF, Orbitrap, FT-ICR, ion trap and quadrupole [76, 95].

A quadrupole is the simplest of mass analysers which is in more advanced MS instruments rather used as a mass filter. The quadrupole mass analyser consists of four, parallel-aligned cylindrical rods (Figure 1.10). Ions passing through the quadrupole are filtered based on their stability in a field of an oscillating radio frequency voltage. The stability of the ions at a given radio frequency voltage ( $V$ ) or direct potential ( $U$ ) can be calculated according Equation 1.4 and Equation 1.3, in which  $a_u$  and  $q_u$  represent the stability space,  $m/z$  is the mass-to-charge ratio,  $r_0$  is the spacing between the rods,  $\omega$  is the angular frequency, and  $e$  is the elementary charge [96]. Since  $\omega$ ,  $r_0$ , and  $e$  are constant, a changing the  $m/z$  is directly proportional to  $a_u$  and  $q_u$ .

$$U = a_u \frac{m}{z} \frac{\omega^2 r_0^2}{8e} \quad (1.3)$$

$$V = q_u \frac{m \omega^2 r_0^2}{z 4e} \quad (1.4)$$



**Figure 1.10:** Schematic representation of a quadrupole mass analyser. Adapted from [96]

In a TOF mass analyzer, ions are separated based on their different flight times in a field-free region. Ions entering the MS are collected and accelerated. The velocity ( $v$ ) is depended on the kinetic energy ( $E_K$ ) and the mass ( $m$ ) (Equation 1.5). Since the ions are given the same kinetic energy at the acceleration, ions with a smaller mass will arrive at the detector faster than ions with a greater mass [96].

$$v = \sqrt{\frac{2E_k}{m}} \quad (1.5)$$

The time ( $t$ ) required to travel distance ( $L$ ) can then be calculated according to Equation 1.6 [96].

$$t = \frac{L}{v} \quad (1.6)$$

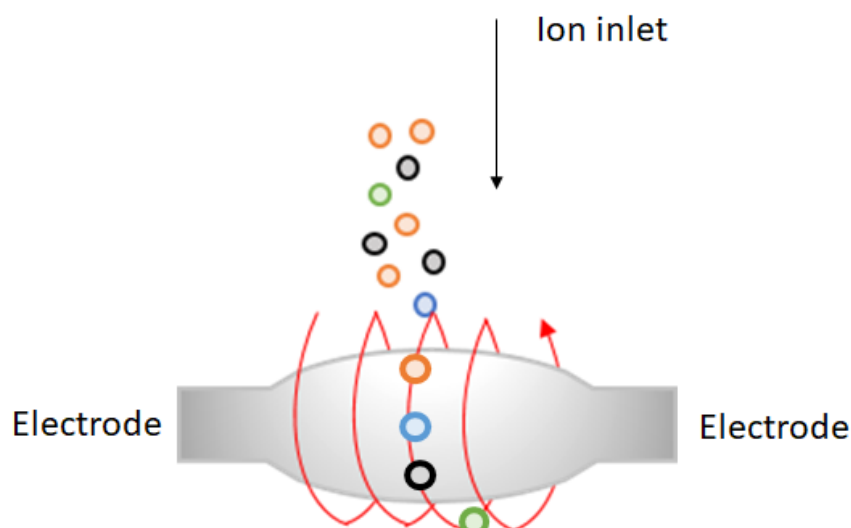
Since  $E_K$  is equal to the product charge ( $z$ ), elementary charge ( $e$ ) and the potential ( $V_s$ ), Equation 1.5 and 1.6 can be rewritten to Equation 1.7 [96]. That shows that  $m/z$  is proportional to  $t^2$ .

$$t^2 = \frac{m}{z} \left( \frac{L^2}{2eV_s} \right) \quad (1.7)$$

An Orbitrap mass analyser uses the oscillation frequency of ions between two electrodes whilst orbiting [96, 97]. A schematic representation of the Orbitrap is shown in Figure 1.11. The oscillation frequency ( $\omega$ ) can be calculated according to Equation 1.8, in which  $k$  is constant based on the curvature of the electric field [97]. Equation 1.8 can be rewritten to Equation 1.9, to show that an ion's  $m/z$  is inversely related to the  $\omega$  squared. In an Orbitrap mass spectrometer, the oscillation frequency is measured over a set scan time and subsequently converted to an  $m/z$  scale via Fourier transformation [96]. Orbitrap instruments are well known for their high scan speed, mass resolving power and mass accuracy, allowing confident identification of molecules of interest [98].

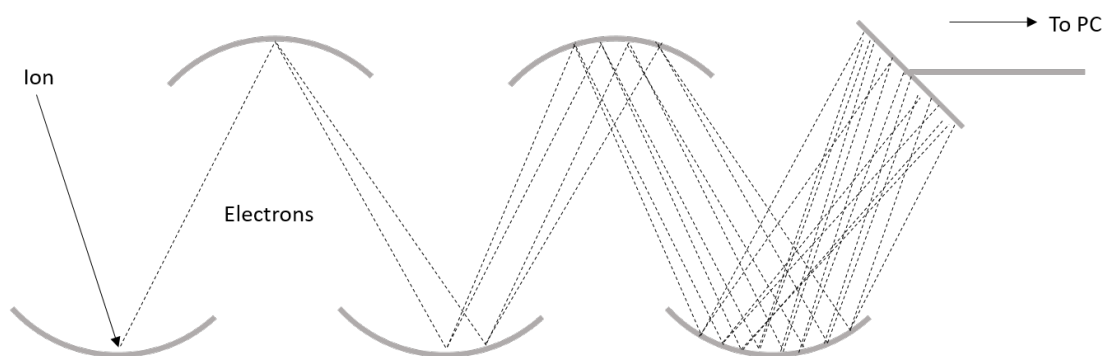
$$\omega = \sqrt{\left(\frac{z}{m}\right)} k \quad (1.8)$$

$$\frac{m}{z} = k \cdot \frac{1}{\omega^2} \quad (1.9)$$



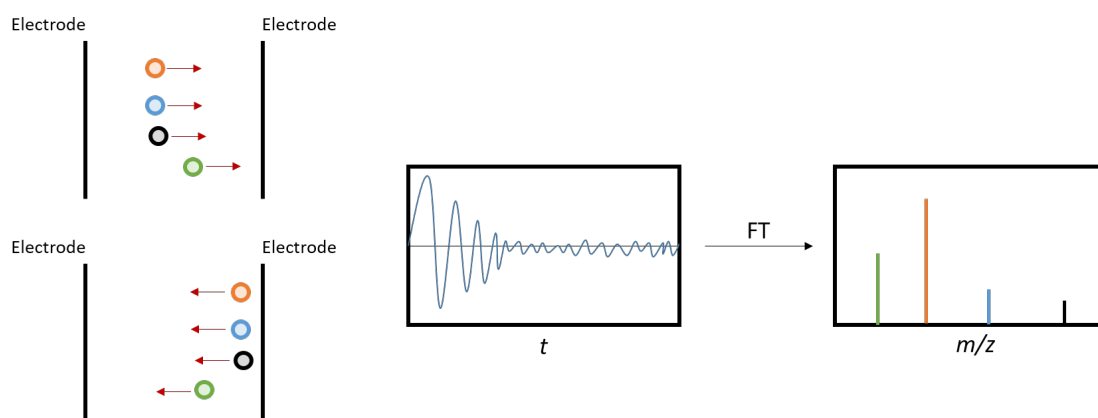
**Figure 1.11:** Schematic representation of the Orbitrap mass spectrometer

**Detection** Nowadays MS instrument use a variety of systems to detect and amplify the signal. Electron multipliers are most commonly used [96]. In this type of the detector the ion hits a dynode which converts the ion into an electron. The electron is then reflected to another dynode which doubles the amount of electrons. This is then repeated  $n$  times before the actual read-out (Figure 1.12). This process can amplify the signal by 4 to 5 orders of magnitude [96].



**Figure 1.12:** Schematic representation of an electron multiplier. Adapted from [96]

For Fourier-transform (FT) instruments, like an Orbitrap, the detection is significantly different. The signal (image current) is created when ions pass near an electrode or detection plate [99]. The image current (time domain) is then converted into the frequency domain via Fourier transformation (Figure 1.13) [96]. The frequency scale is then calibrated to convert it to a  $m/z$  scale (mass spectrum) [96].



**Figure 1.13:** Schematic representation of detection in a FT instrument

### 1.6.1.3 IDENTIFICATION OF PROTEINS VIA MASS SPECTROMETRY

Prior to the availability of computational tools for peptide sequencing, annotation had to be done manually. With the increase of MS for proteomics and the introduction of shotgun proteomics (untargeted analysis) [100], automated identification of peptides from MS/MS spectra became a necessity. Throughout the years, several identification algorithms have been developed (Table 1.4).

**Table 1.4:** Available search algorithms for peptide identification from MS/MS spectra

Algorithm	Introduction	Reference	Open source
SEQUEST	1994	[101]	No
MASCOT	1999	[102]	No
OMSSA	2004	[103]	Yes
X!Tandem	2004	[104]	Yes
OMSSA	2004	[103]	Yes
Andromeda	2011	[105]	Yes
Tide	2011	[106]	Yes
Comet	2013	[107]	Yes
MS-GF+	2014	[108]	Yes
MSAmanda	2014	[109]	Yes

Each algorithm has its own method of scoring MS/MS spectra, which makes that different algorithms will give the user provide different results [110, 111]. However, the concept of identifying peptides from MS/MS spectra in each algorithm is similar. The user will provide with a list of proteins in .FASTA format (Figure 1.14). The search algorithm will then generate theoretical peptides based on the used enzyme and the number of allowed missed cleavages. A missed cleavage means that the enzyme did not cut at the expected location. This is most often by inaccessibility of the cleavage site [112]. The search algorithm will then compare the theoretical spectra to the acquired spectra based on user-defined mass tolerance window for the intact peptide ion and the fragment ions. It will then generate a score based on correlation or probability.



## CHAPTER 1

To validate spectra, a list of decoy peptides is used which is in general each protein sequence in the .FASTA file reversed [113]. These decoy peptides are then compared to the acquired MS/MS spectra. Since the decoys are non-existing peptide sequences, MS/MS spectra matching a decoy peptide can be considered a false discovery. The user has to define a false discovery rate (FDR) (in general 1%) as cut-off to define which peptides can be considered identified [114].

As a complementary identification strategy to sequence database searching, the use of spectral libraries has become a more popular approach [115, 116]. Spectral libraries are commonly used for identification of proteins from data-independent acquisition (DIA) data sets through software tools such as Skyline [117], Spectronaut and DIA-Umpire [118]. A library is built by a number of data-dependent acquisition (DDA) files containing the fragmentation spectra of the target proteins identified through sequence database search. Alternatively, new tools have emerged to predict fragmentation of peptide through artificial neural networks and build subsequent spectral libraries based on *in silico* fragmentation spectra [119–122].

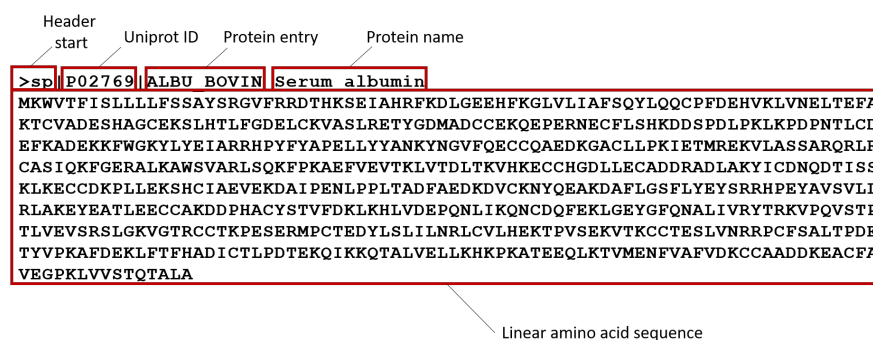
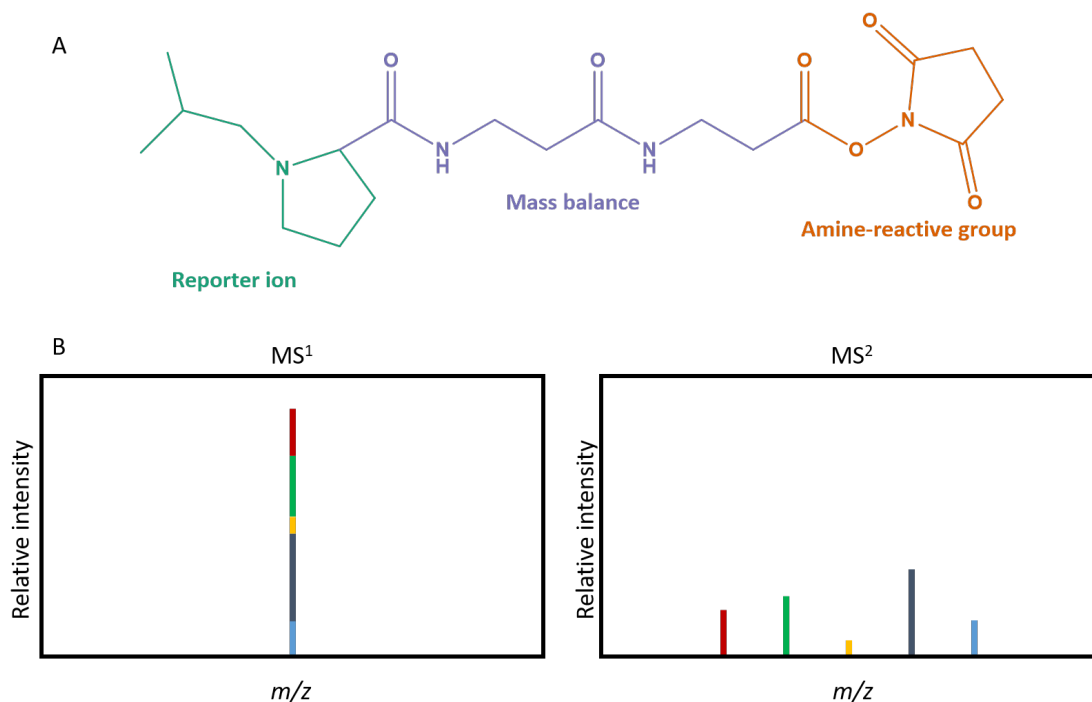


Figure 1.14: The .FASTA format explained

### 1.6.1.4 QUANTIFICATION OF PROTEINS VIA MASS SPECTROMETRY-BASED STRATEGIES

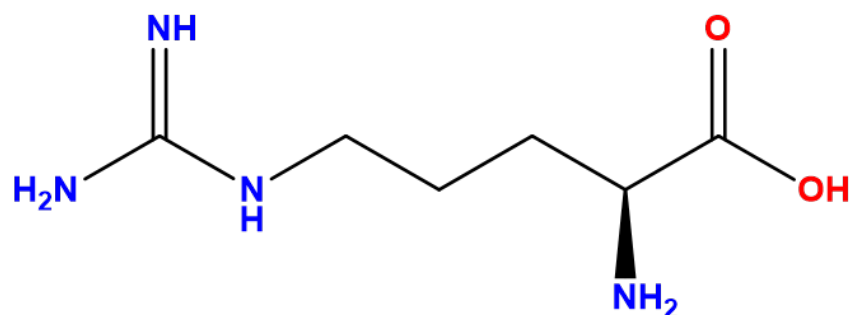
**Tandem mass tags (TMT)** In TMT labeling, peptides between different biological conditions are labelled with different chemical tags. Commercially available tags (Thermo Fisher) consists of a maximum 16 different tags. Thus, 16 different biological conditions can be compared to each other. At,  $MS^1$  level, all tags will have the same precursor ion.

However, during fragmentation all tags will lose their reporter ion, which is different in mass and therefore will have a different  $m/z$  value (Figure 1.15). Using the intensities of the reporter ions, relative quantification can be performed [123].

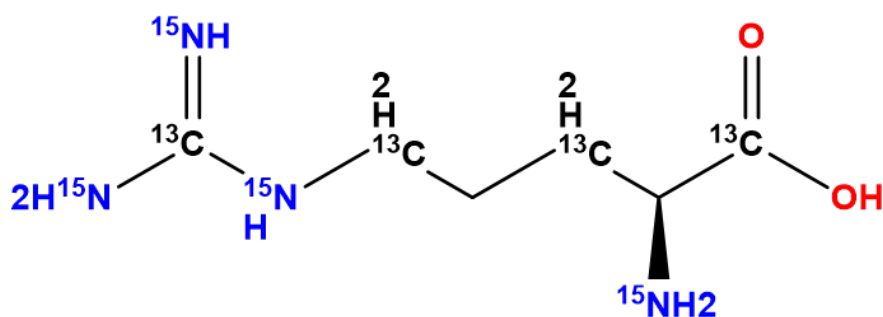


**Figure 1.15:** The use of TMT for relative quantification of proteins. (A) Example of a tandem mass tag. The amine-reactive group is used for covalently bind to the peptide. Each tag has a different reporter ion which make the tags distinguishable at  $MS^2$ . The balance part is essential to make up the mass difference between reporter ions to generate the same precursor ion. (B) During fragmentation the bond between the reporter ion and balance part is cleaved, permitting relative quantification based on the intensity of the reporter ion. Adapted from [124].

**Stable isotope labelling with amino acids in cell culture (SILAC)** A SILAC experiment consists of comparing two conditions by growing cells or feeding subjects in-/with a either medium containing a light label or a heavy label. Because trypsin is commonly used as an enzyme for MS-based proteomics, the label of choice is mostly a heavy arginine or lysine (Figure 1.16). These 'heavy' amino acids are  $^{13}C$ ,  $^{15}N$ -labelled in order to distinguish them from any natural occurring interference. The signal intensities at  $MS^1$  are then used for relative comparison of the protein abundance [125, 126].



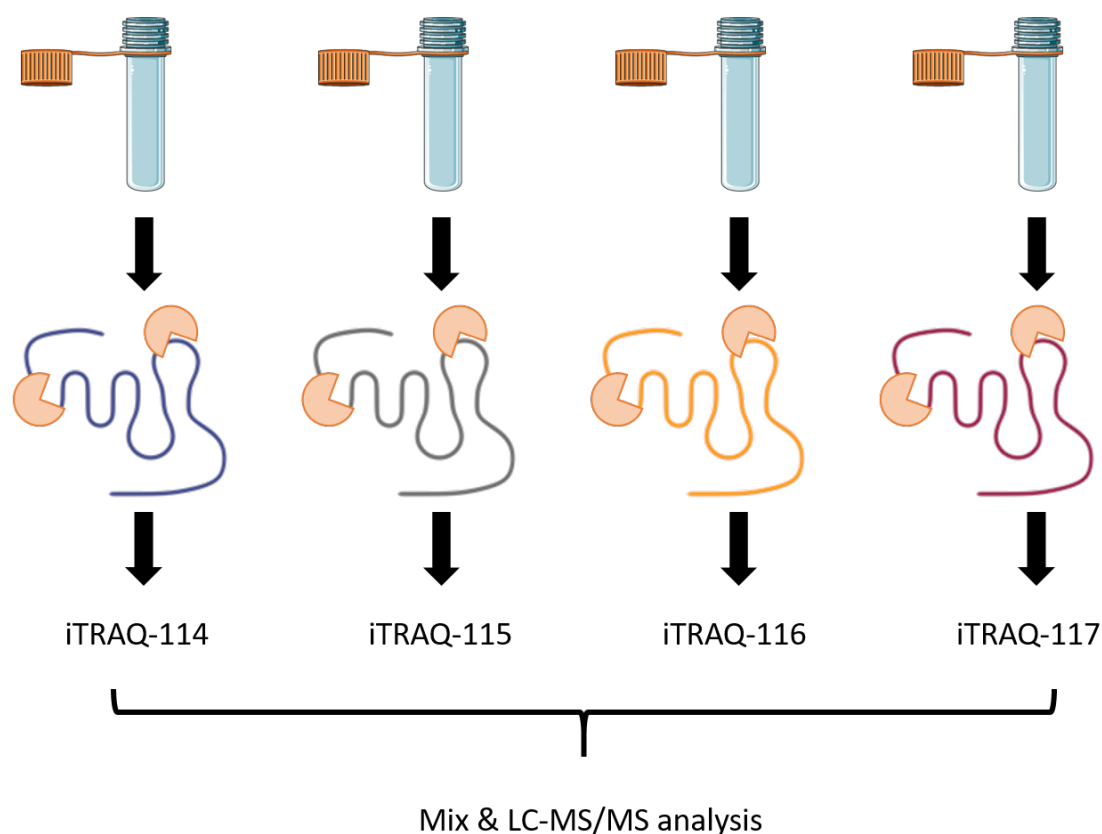
**Exact Mass: 174.11**



**Exact Mass: 182.11**

**Figure 1.16:** 'Light' (top) and 'heavy' arginine (bottom)

**Isobaric tag for relative and absolute quantification (iTRAQ)** ITRAQ can be used for absolute and relative quantification of proteins. In this strategy, isotopically labelled tags are used like with TMT [127, 128]. In comparison to TMT, a maximum of four different conditions (4-plex) can be assessed at a time. In contrast to TMT, the chemical tags are all different in mass and will therefore have a different  $m/z$  value at  $MS^1$  level (Figure 1.17). The intensities from the precursor ions can then be used for relative and absolute quantification [128, 129]. Absolute quantification can be achieved by using the isobaric tags as internal or external calibration standards [130]. However, iTRAQ is predominantly used for relative difference in protein expression [131].



**Figure 1.17:** Workflow for iTRAQ-based quantification of proteins

**Label-free quantification** As opposed to the above mentioned strategies, relative quantification of protein abundance can also be performed without any labelling. A number of reasons to choose for label-free quantification are reduced expenses (no expensive reagents required), no additional sample preparation is required, and it is suitable for discovery experiments (shotgun proteomics) [132]. For quantification,  $MS^1$  intensities,  $MS^2$  intensities or a combination of both are used as a measure of relative quantity. In order to get reliable quantification, the data needs to be normalised to account for variance induced by technical sources e.g. instrument response and digestion efficiency [133–136].

### 1.6.1.5 MS ACQUISITION STRATEGIES FOR PROTEOMICS

Classically, discovery-driven proteomics have used DDA of mass spectrometry data. In DDA, selection of ions for fragmentation is done based on their intensity [76]. Ions which have an intensity higher than a user-defined threshold are sent to the HCD cell for fragmentation and a MS/MS spectrum is acquired subsequently. However, intensity-based selection could lead to underrepresentation of low-abundant peptide or protein ions. As alternative strategy, DIA was introduced [137]. In DIA, the MS does not isolate single precursor ions, but instead large  $m/z$  windows ( $> 20 m/z$ ) are selected and all ions present in the  $m/z$  range are fragmented at once.

### 1.6.2 LIMITATIONS OF STANDARD MASS SPECTROMETRY-BASED APPROACHES FOR HIGH THROUGHPUT ARRAY SCREENING

Current LC-MS/MS strategies provide excellent in-depth, quantitative proteomics data. However, as the standard workflow is laborious and time-consuming due to extraction, several purification and concentrations steps, protein digestion and long LC-MS/MS runs ( $> 1$  hour) per sample. For screening of a vast number of samples, e.g. a material microarray ( $\sim 500$ ), weeks of labour and instrumentation time are required to complete the analysis. Therefore, a more automated approach with minimal sample preparation is required to make the analysis compatible with array screening.

## 1.7 AMBIENT MASS SPECTROMETRY FOR HIGH THROUGHPUT SCREENING

### 1.7.1 THE DEFINITION OF AMBIENT MASS SPECTROMETRY

In ambient MS, ions are generated under ambient conditions [138, 139]. Further advantages of ambient MS are the ability to perform MS without upfront chromatography allowing increased sample throughput (reduced run time per sample) and reduced sample preparation [140, 141]. Ambient MS is already an established analysis strategy

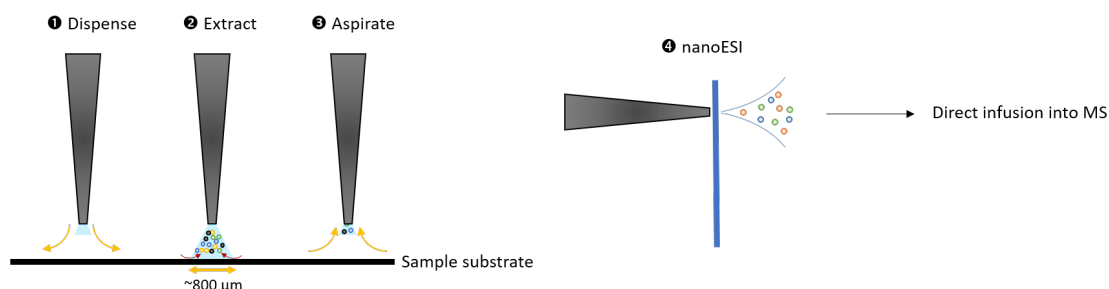
for drug discovery, clinical, environmental and forensic applications [142–147].

### 1.7.2 APPLICATIONS OF AMBIENT MASS SPECTROMETRY IN PROTEOMICS

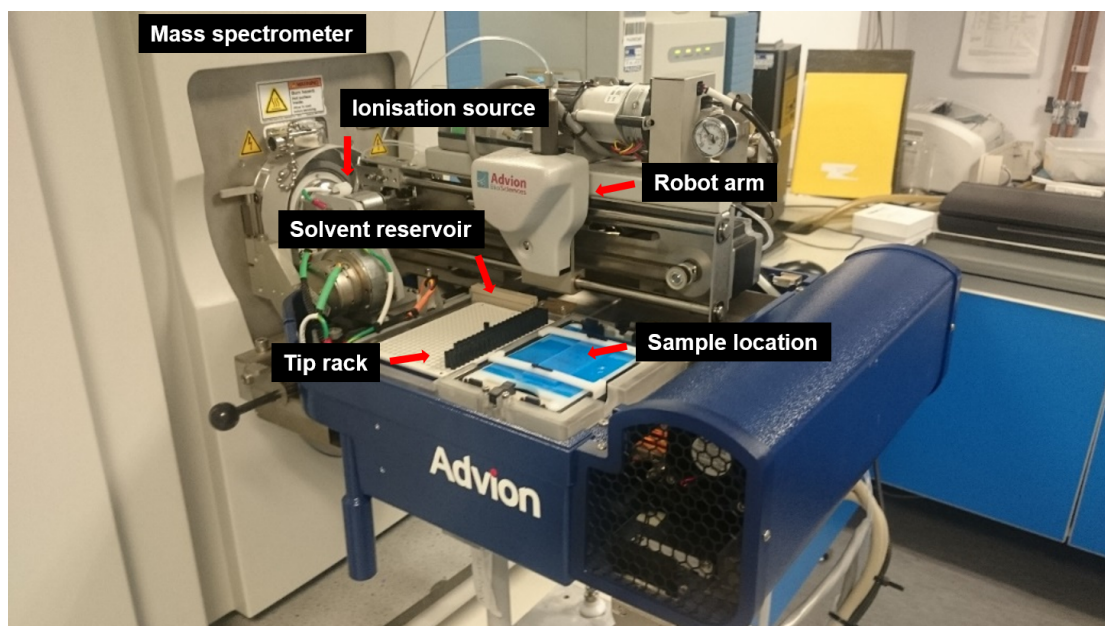
Several studies have focused on the identification of proteins using ambient ionisation techniques. Montowska *et al.* used liquid extraction surface analysis (LESA) for studying peptide markers as tool for meat authentication [148–150]. Rao *et al.* [151] used desorption electrospray ionisation (DESI) [152] and LESA [153] for the analysis of proteins deposited on biomaterial surfaces. Other proteomics applications of LESA were reported for intact protein analysis in bacterial colonies [154, 155] and dried blood spot analysis [156–160].

## 1.8 LESA-MS/MS AS A TOOL FOR HIGH THROUGHPUT QUANTITATIVE SURFACE PROTEOMICS

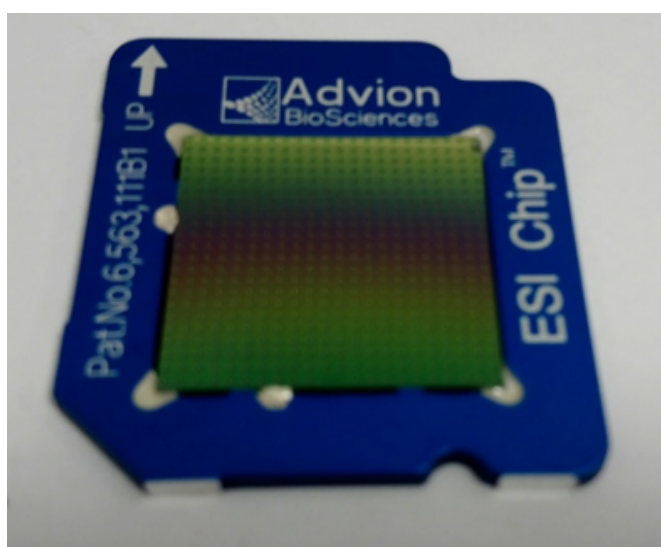
LESA (Figure 1.18) is an ambient MS technique invented by Kertesz & Van Berkel [153]. The system consists of a fully-automated robotic system to dispense and aspirate a user-defined extraction solvent. The system (Figure 1.19) uses a single disposable conductive pipette tip per analysis. Once the dispensing/aspiration cycle has been completed, the tip containing the extract is directed towards a chip (Figure 1.20) at the front end of the instrument which serves as an ionisation source.



**Figure 1.18:** Schematic overview of the principle behind LESA



**Figure 1.19:** The commercial available LESA system (TriVersa Nanomate) from Advion Biosciences (Ithaca, NY)



**Figure 1.20:** Chip used for nanoESI in LESA-MS/MS analysis

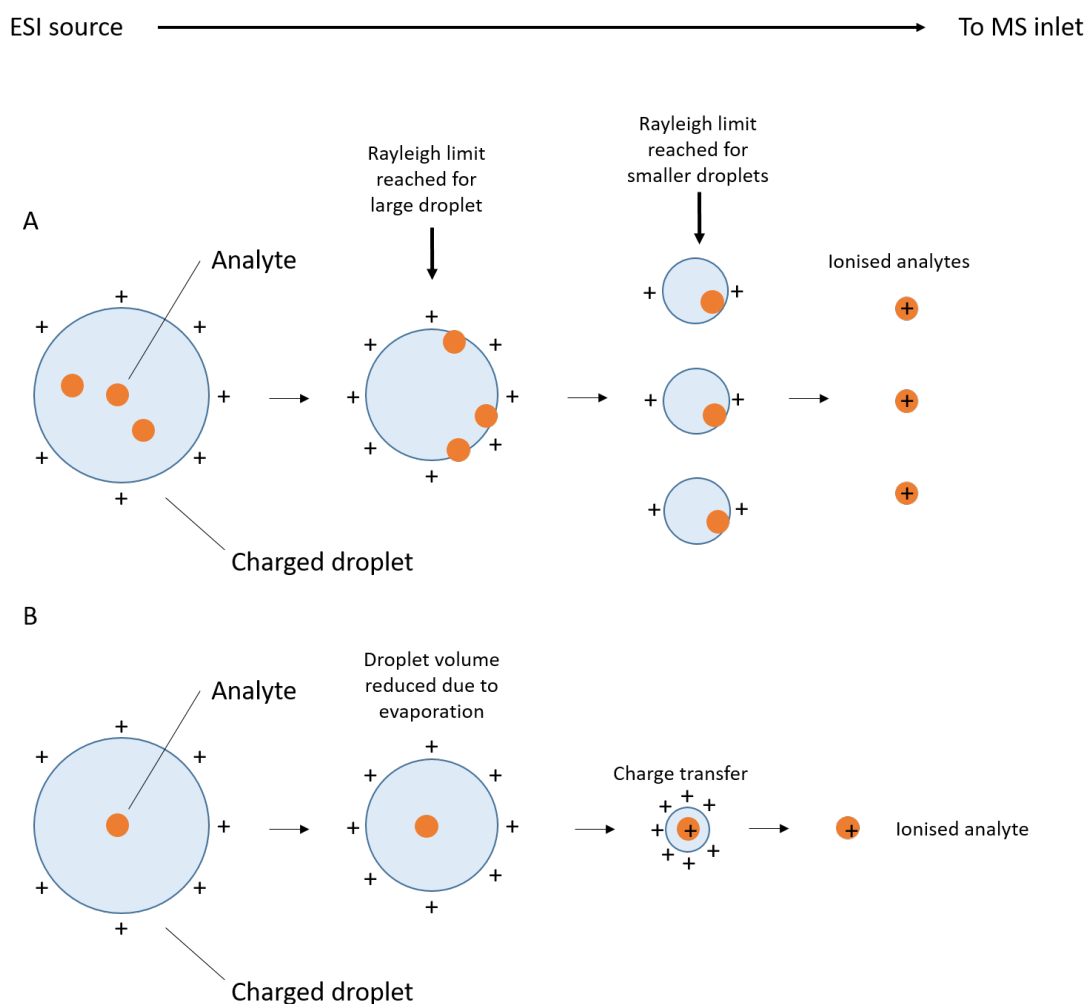
The solvent containing the extracted analytes is ionised via nano-electrospray ionisation (nanoESI) [161], which is similar to electrospray ionisation (ESI) [162, 163] but lower flow rates are used, hence, increasing the sensitivity [161, 164]. The principle behind ESI is not fully understood, but two main theories exist on the ionisation process

[165, 166]. The first theory is the *ion evaporation model* [167] (Figure 1.21A), which postulates that ions are generated by coulombic repulsion when the surface charge density of a droplet reaches the Rayleigh limit [168]. The Rayleigh limit is defined as follows (Equation 1.10) in which  $q$  is the charge,  $\gamma$  is the surface tension,  $\epsilon_0$  and  $r$  is the radius of the droplet. Once the Rayleigh limit is reached or nearly reached, multiple smaller charged droplets are generated. This process continues until the the charged droplet arrives at the MS [166].

$$q^2 = 64\pi\gamma\epsilon_0r^3 \quad (1.10)$$

The other theory is the *charge residue model*. This theory states that each droplet formed via ESI contains only one analyte ion [166]. The analyte ion is then released when the solvent has completely evaporated [166] (Figure 1.21B).





**Figure 1.21:** Schematic representations of ionisation theories for ESI. (A) The *ion evaporation model*. The droplet evaporates until the charge density reaches the Rayleigh limit (B) The *charge residue model*. Every droplet contains a single analyte from which the ion is released once the solvent has evaporated. Adapted from [169]

LESA shows great potential to be used for screening of material microarrays due to the fully automated extraction and MS analysis. Further, the technical variance is reduced due to the minimal amount of sample preparation required for ambient MS and minimal manual handling of samples [170]. Further several studies (e.g. [84, 148–151, 156–160, 171–174]) have shown that LESA is a suitable analysis platform for proteomics. The group of Cooper has extensively used LESA for studying proteins and protein complexes in their native state from several biological substrates [84, 156–158, 160, 174],

have pushed LESA-MS/MS analysis to detect larger protein complexes (>70 kDa) [175], have incorporated different ion mobility techniques to obtain conformational information on intact proteins [176–180] and have explored to use of LESA for imaging native proteins in tissue section [177, 181, 182].

Montowska *et al.* have successfully used LESA to identify peptide markers for authentication of different meat species [148–150]. Proteins were either in solution or *in situ* digested and analysed via LESA-MS/MS.

Rao *et al.* were the first to explore the use of LESA-MS/MS for the analysis proteins adsorbed onto biomaterials surfaces [151]. Protein standards (individual and mixtures) were deposited on Permanox<sup>TM</sup> cell culture slides, *in situ* digested and analysed using LESA-MS/MS. They found that proteins could be detected down to concentrations of pmol·mm<sup>-1</sup>. The next step is now to further explore this methodology on proteins relevant to cell culture media.

## 1.9 AIMS & OBJECTIVES THESIS

The overall aim of this thesis will be to develop a strategy for quantitative MS-based proteomics analysis which is compatible with material microarray screening. Being able to assess quantitatively measure protein adsorption on combinatorial libraries of synthetic polymers will permit understanding of cell-material as well as protein-material interactions. The basic mechanistic understanding of cell response on synthetic polymers will allow rationalising as well as improvement of biomaterial design.

The first aim of the chapter will be to define a LESA-compatible substrate to be used for material microarray analysis (Chapter 2). Next, all protein digestion, LESA and MS parameters need to be defined for performing quantitative surface analysis of digested proteins (Chapter 3). Then, the method can be extended for analysis of synthetic polymers in array format (Chapter 4) as well as the analysis of scaled-up polymers in 6-well plate format (Chapter 5) in order to gain more understanding on the cell-instructive mechanism of synthetic polymers. Further, the use of LESA will be

## CHAPTER 1

extended for untargeted metabolomics of clinical arrays (tumour microarrays) in order to obtain valuable clinical information from minimal sample size (Chapter 6).

# Improving LESA-MS sampling reproducibility through a superhydrophilic-superhydrophobic substrate<sup>1</sup>

## 2.1 INTRODUCTION

**F**OR quantitative MS analysis, the variance induced from technical sources should be minimised to be able to reliably measure biological differences. Examples of sources of technical variation are the extraction protocol, the effect of reagents, and data acquisition strategy [183]. In LESA-MS, the main sources of technical variation is the poor repetitive sampling due to spreading of the extraction solvent on the surface of interest [184]. Irregular spreading of the extraction solvent could potentially lead to poor reproducibility of the re-aspirated solvent volume [184] and therefore may introduce undesired technical variance between replicate samples. Brenton & Godfrey [185] limited the spreading of the extraction solvent by using electrodes. However, they developed an in-house system which is not commercially available to date. Furthermore,

---

<sup>1</sup>This chapter has been fully published as technical note: J. Meurs *et al.*, *Anal. Chem.* **90**, 6001–6005 (2018). doi: 10.1021/acs.analchem.8b00973

the sample probe needs to be washed after every single analysis to avoid carry-over effects. Almeida *et al.* [184] modified the gas supply line of the commercially available TriVersa Nanomate, which can perform liquid extraction surface analysis. By modifying the gas line, the extraction solvent could be pressurised and therefore limiting the solvent spread. Other methods to limit the solvent spread are bringing the tip with extraction solvent directly into contact with the surface ("contact" LESA [155]) or by ablating analytes from the surface using a laser and capture the volatile analytes in a droplet of extraction solvent [186].

It has been noted that the spreading of the extraction solvent on hydrophilic surfaces is significant [155] resulting in limited solvent recovery. Furthermore, the use of non-polar solvents such as chloroform dramatically increases the solvent spread [184]. When the surface itself is not part of the system and is simply there to support a spot of liquid analyte, the analytical performance of liquid surface sampling could be enhanced by the judicious choice of the surface properties to vastly improve spatial confinement of both sample deposition and the application of extraction solvent. Van Berkel *et al.* reported a reproducibility ranging between 10% and 34% for different lysozyme concentrations on Teflon-masked microscope slides analyzed by a continuous surface sampling probe [187].

Hence, repeatable extraction solvent recovery as well as extraction efficiency could be optimised by designing a substrate surface, which confines both the liquid sample and the LESA extraction solvent within a consistent area. A platform which meet those conditions is the Droplet Microarray (DMA). The DMA consists of a pattern of superhydrophilic spots bordered by a superhydrophobic material produced by chemically modifying a polymer surface [188, 189]. These surface properties are of interest for LESA to assist in obtaining reproducible analyte spot deposition and extraction solvent coverage of the spots. The superhydrophilic part guides the solvent along the surface while the superhydrophobic material has a low surface tension and is therefore difficult

to wet [190]. This will limit the deposited sample as well as the extraction solvent from further spreading. Besides the surface chemistry, the use of the DMA for LESA is of interest due to its high-throughput design. In this chapter, the DMA high-throughput array is compared to a standard glass microscope slide, which is often used as a substrate for LESA-MS [84, 139, 171, 172, 184, 186, 191–198], to test the hypothesis that confining the sample as well as the extraction solvent in a defined area assists in obtaining repeatable solvent recovery leading to improved spectral reproducibility and multivariate modeling. A mixture containing a number of components is used to test the extraction repeatability. The spectral quality and biological interpretation is assessed by performing an untargeted metabolomics experiment using readily available urine samples.

## 2.2 MATERIALS & METHODS

### 2.2.1 SAMPLE PREPARATION

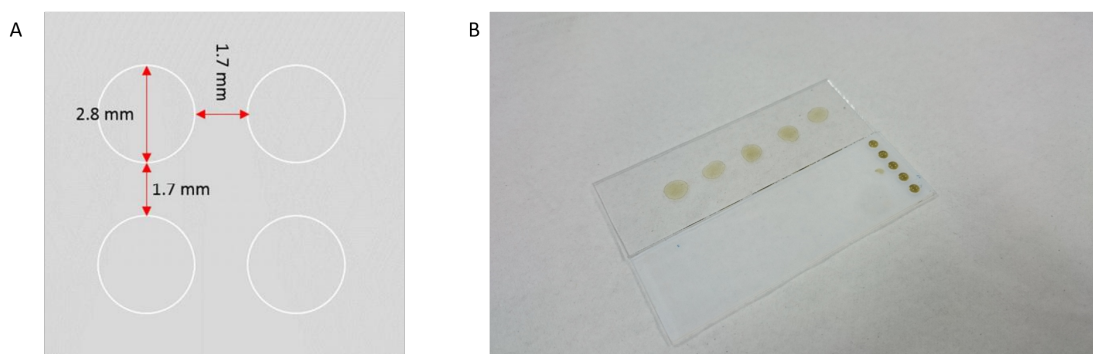
#### 2.2.1.1 STANDARD MIXTURE

The components of the standard mixture can be found in Table 2.1. Chemicals were chosen to cover a wide range of molecular weights and physicochemical properties. Stock solutions for taurine, L-arginine, diphenhydramine HCl, Rhodamine 6G, raffinose pentahydrate (all 10 mM), and vitamin B12 were prepared in deionised water (18.2 MO; Elga PureLab, Lane End, High Wycombe, UK). A stock solution for hemin was prepared in 50 mM sodium hydroxide (98.5–100.5%; VWR International, Leuven, Belgium). Subsequently, all stock solutions were mixed and diluted with deionised water to a final concentration of 10  $\mu$ M for each standard.

**Table 2.1:** Components of the standard mixture

Component	Supplier	Purity
Taurine	Acros Organics	99%
L-arginine	Sigma-Aldrich	$\geq 98\%$
Diphenhydramine HCl	Sigma-Aldrich	98%
Rhodamine 6G	Acros Organics	99%
Raffinose pentahydrate	Sigma-Aldrich	98%
Hemin	Sigma-Aldrich	$\geq 97\%$
Vitamin B12	Sigma-Aldrich	$\geq 98\%$

The standard mixture was manually pipetted onto 10 individual features of a superhydrophobic-superhydrophilic array (Droplet Microarray (DMA), Aquarray, Karlsruhe, Germany) [188] and dried in a fume hood ( $\sim 1$  hour). The DMA consists of a pattern of 2.8 mm circular superhydrophilic spots bordered by 1.7 mm superhydrophobic material (Figure 2.1). Further, a  $5 \times 2$  grid was created on a microscope glass slide (75 mm  $\times$  25 mm) to deposit the standard mixture on. The glass slide was also dried in a fume hood ( $\sim 1$  hour) prior to further analysis.



**Figure 2.1:** (A) Schematic representation of the dimensions of a spot on the Droplet Microarray. Superhydrophilic spot size: 2.8 mm; superhydrophobic border width: 1.7 mm (B) Macroscopic image of a 10 mM hemin solution dispensed ( $V = 3 \mu\text{L}$ ) onto glass (top) and DMA (bottom). Images reveals the trapping of the aqueous solution within the superhydrophobic areas whilst the solution spreads out widely on a microscope glass slide

## CHAPTER 2

### 2.2.1.2 URINE SAMPLES

Urine samples from an intervention study were provided by Dr Sergey Evseev (School of Pharmacy, University of Nottingham). In this study, four healthy male volunteers participated. Each volunteer had to provide a urine sample prior to intervention (control). The intervention consisted of drinking a cup of Earl Gray tea ( $\sim 200$  mL). Volunteers had to provide a urine sample within two hours of intervention. After collection, samples were immediately stored at  $-80^{\circ}$  C. The study design was authorized by University of Nottingham Ethical Committee of the School of Pharmacy, University of Nottingham (reference number: 021-2016).

For the repeatability assessment, 5 replicate spots ( $V = 2$   $\mu$ L) were dispensed on both surfaces using the urine sample of one subject. The urine spots were allowed to dry in a fume hood ( $\sim 1$  hour) and analysed with LESA-MS. For the untargeted metabolomics experiment (before vs. after tea intervention), 2  $\mu$ L of urine was dispensed onto both glass and DMA for each subject for each condition. The samples were allowed to dry in a fume hood ( $\sim 1$  hour) and analysed with LESA-MS.

### 2.2.2 LIQUID EXTRACTION SURFACE ANALYSIS-MASS SPECTROMETRY

LESA (TriVersa Nanomate, Advion Biosciences, Harlow, UK) was performed on each sample spot using an extraction solvent containing LC-MS grade methanol (MeOH; CHROMASOLV; Riedel-Haen, Seelze, Germany), LC-MS grade water ( $H_2O$ ; CHROMA SOLV; Riedel-Haen, Seelze, Germany) and LC-MS grade formic acid (FA; Optima<sup>TM</sup>, Fisher Scientific) (70:30:0.1 v/v/v) [191]. The total extraction solvent volume was set to 5  $\mu$ L. For 5 seconds, 3  $\mu$ L of extraction solvent was deposited on a sample location. Thereafter, 3.5  $\mu$ L was re-aspirated and hold for 10 seconds. After, the tip was directed towards the nanoESI chip (ESI Chip<sup>TM</sup>, Advion Biosciences, Ithaca, NY). Ionisation was performed at a voltage of 1.45 kV and 0.3 psi gas back pressure [191]. Data were acquired in positive ionisation mode on an Exactive Orbitrap mass spectrometer (Thermo Scientific, Hemel Hempstead, UK). The scan range, resolution and maximum



## CHAPTER 2

injection time were set to  $m/z$  100-1500, 100,000 (at  $m/z$  400) and 1000 ms respectively. The AGC target was set to  $1 \times 10^6$  and the capillary temperature was set to 250°C.

### 2.2.3 DATA ANALYSIS

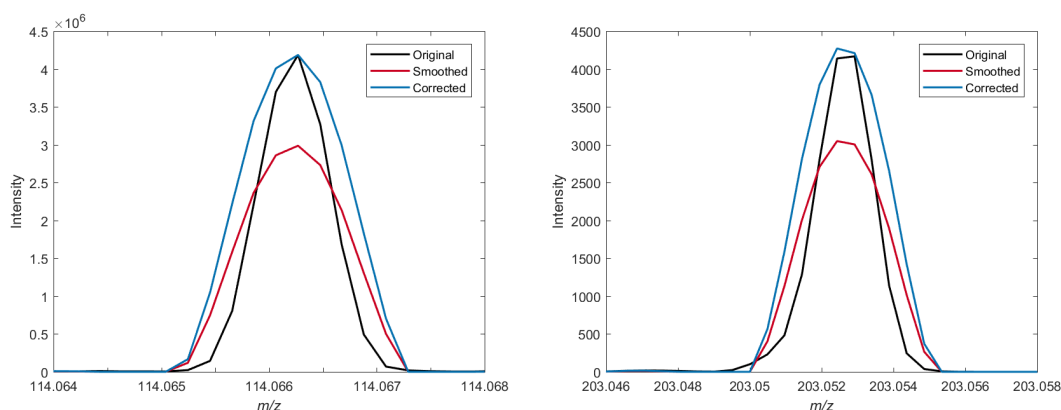
#### 2.2.3.1 SAMPLING REPEATABILITY

Raw files were averaged in Xcalibur v2.1 (Thermo Scientific, Hemel Hempstead, UK) and converted into a .mzXML format using ProteoWizard (v3.0.18178 64-bit) [199]. The obtained files were subsequently parsed into MATLAB (R2016b, The MathWorks Inc.) and peak lists were generated per file. The most intense adduct for each standard mixture component was used to compare relative standard deviations between surfaces using Forkman's F-test [200]. Peak heights were used as a measure of ion intensity.

#### 2.2.3.2 UNTARGETED METABOLOMICS

MS data from urine files were converted to .mzXML as described above and parsed into MATLAB. Zero-filling for Orbitrap data was performed to create a common  $m/z$  vector [201]. Spectra were smoothed using the Savitzky-Golay algorithm [202]. Peak picking was performed using the `findpeaks` function in MATLAB. Subsequently, peak intensities (peak height) were corrected for loss in signal during the smoothing process according to Equation 2.1 in which  $I_{corr,x}$  is the corrected signal intensity for a given peak,  $I_{s,x}$  is the peak intensity after smoothing,  $I_{0,max}$  is the base peak intensity prior to smoothing, and  $I_{s,max}$  is the base peak intensity after smoothing. An example of signal correction is shown in Figure 2.2

$$I_{corr,x} = I_{s,x} \frac{I_{0,max}}{I_{s,max}} \quad (2.1)$$



**Figure 2.2:** Signal intensity before Savitzky-Golay smoothing (black), after Savitzky-Golay smoothing (red) and after signal correction; Left: base peak (creatinine;  $m/z$  114.0663); Right: peak at  $m/z$  203.0529 ( $C_6H_{12}O [M+Na]^+$ ). It can be seen that after correction the signal intensity is restored at its original level

Peaks were matched within a 5-ppm  $m/z$  window and corrected intensities per peak were retrieved. The intensity matrix was used for PCA. Prior to PCA, peak intensities were normalised to the total ion current (TIC) and Pareto scaled [203]. Peaks with a detection rate lower than 80% across all samples or an intensity smaller than 1,000 were excluded [204] for PCA. Missing values were replaced by using the average intensity for corresponding ions (mean value imputation) [205]. Individual peak intensities were statistically analysed using Student's  $t$ -test. Statistical significance was corrected using the Benjamini-Hochberg FDR procedure [206].

## 2.3 RESULTS & DISCUSSION

### 2.3.1 ROBUST LESA WITH DROPLET MICROARRAY

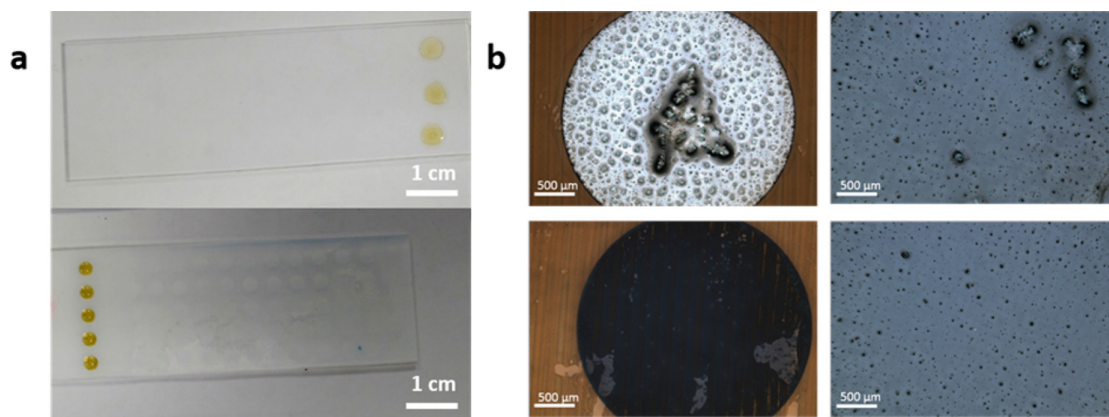
#### 2.3.1.1 STANDARD MIXTURE

Ten spots of 2  $\mu$ L standard mixture were created on glass and DMA and dried under ambient conditions. The standard mixture contains components covering the range  $m/z$  100-1500 and would therefore be representative for a metabolomics experiment. Furthermore, the used standards have different physicochemical properties and there-

fore potentially different LESA extraction efficiency. Spots were analysed by LESA-MS using an extraction solvent composed of MeOH, H<sub>2</sub>O, and FA (70:30:0.1 v/v/v). The most intense adduct for each component was used for statistical comparison. From Table 2.2 can be derived that all relative standard deviations (RSDs) are significantly lower (Forkman's *F*-test) when using the DMA as sample substrate. This indicates that sampling variation can be reduced by confining the sample as well as the extraction solvent within a defined area (Figure 2.3). Moreover, all RSDs for the DMA substrate were found to be smaller than 15%. A RSD of 15% or smaller is required for quantitative analysis [207].

**Table 2.2:** Reproducibility of signal intensities for every standard mixture component extracted from glass and Droplet Microarray

Component	Adduct	<i>m/z</i>	RSD glass	RSD DMA	p-value
Taurine	[M+Na] <sup>+</sup>	148.0044	25.3%	7.6%	0.015
L-arginine	[M+Na] <sup>+</sup>	197.1015	23.1%	7.3%	0.020
Diphenhydramine	[M+H] <sup>+</sup>	256.1701	20.1%	5.1%	4.0 × 10 <sup>-4</sup>
Rhodamine 6G	[M-Cl] <sup>+</sup>	443.2335	20.0%	5.1%	4.3 × 10 <sup>-4</sup>
Raffinose	[M+Na] <sup>+</sup>	527.1588	21.7%	7.4%	0.0034
Hemin	[M-Cl] <sup>+</sup>	616.1773	21.4%	5.6%	7.7 × 10 <sup>-4</sup>
Vitamin B12	[M+2Na] <sup>2+</sup>	700.2735	26.6%	7.9%	0.0056



**Figure 2.3:** Depositing and extraction of samples from the Droplet Microarray and glass surface. (A) 10 mM hemin in 50 mM NaOH manually pipetted onto glass (top) and Droplet Microarray (bottom). Sample volume: 3  $\mu$ L; spot size DMA:  $\sim$ 2.8 mm; spot size glass:  $\sim$ 4.0 mm. (B) Microscope image (Optical Profile, Zeta Instruments; 5 $\times$  magnification) of dried urine on Droplet Microarray (left) and glass (right). Images taken before (top) and after extraction (bottom). Visually higher efficient extraction from DMA as can be derived from the empty spot in the microscope image. Sample volume: 2  $\mu$ L.

### 2.3.1.2 URINE SAMPLE

The readily available urine samples from a previous study were used "as is". One control urine sample was used to investigate the sampling reproducibility of a complex biological matrix. Two  $\mu$ L urine was dispensed onto five clean spots of the Droplet Microarray and on five marked spots on a clean microscope glass slide. The spots were allowed to dry in a fume hood ( $\sim$  1 hour). After, LESA-MS analysis was performed using MeOH/H<sub>2</sub>O/FA 70:30:0.1 v/v/v as extraction solvent. Peaks were included when they were present in each replicate from both surfaces and their intensities were greater than  $1 \times 10^3$ . Table 2.3 shows the RSDs as well as the p-value after comparison of both substrates using Forkman's F-test.

**Table 2.3:** Statistical comparison of LESA extraction reproducibility from glass and DMA

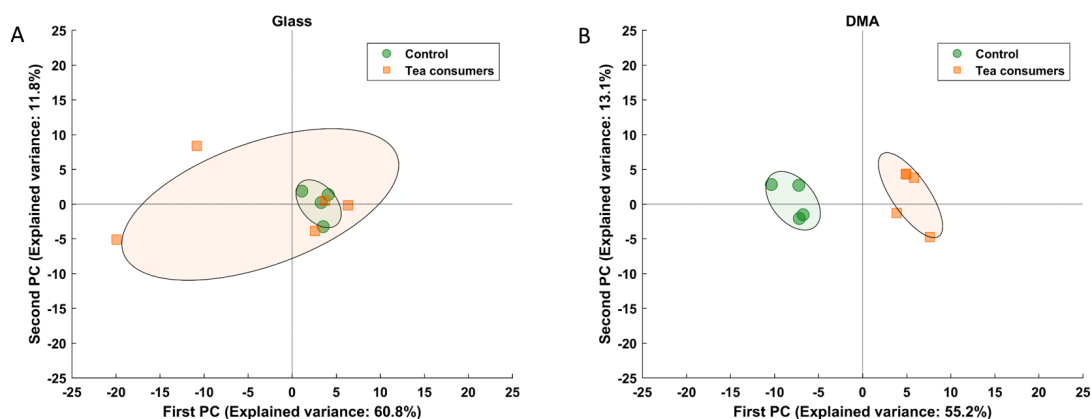
Ion ( <i>m/z</i> )	RSD glass (%)	RSD DMA (%)	p-value
112.8959	93.7	5.5	$8.12 \times 10^{-4}$
114.0664	93.2	1.4	$8.83 \times 10^{-5}$
142.9482	56.2	4.2	$1.57 \times 10^{-3}$
158.9222	91.8	5.2	$6.58 \times 10^{-4}$
166.0725	91.9	5.1	$7.64 \times 10^{-4}$
173.0214	91.8	2.9	$2.57 \times 10^{-4}$
173.0576	17.4	4.3	0.088
180.9040	91.5	5.9	$5.63 \times 10^{-4}$
245.1155	91.8	4.3	$5.33 \times 10^{-4}$
255.0630	91.8	2.8	$2.35 \times 10^{-4}$
283.2038	91.8	2.5	$2.01 \times 10^{-4}$
364.9646	91.8	2.9	$2.39 \times 10^{-4}$
368.9591	92.0	2.9	$2.49 \times 10^{-4}$
483.0767	91.8	3.4	$3.29 \times 10^{-4}$

## 2.4 IMPROVED UNTARGETED METABOLOMICS WITH DROPLET MICROARRAY

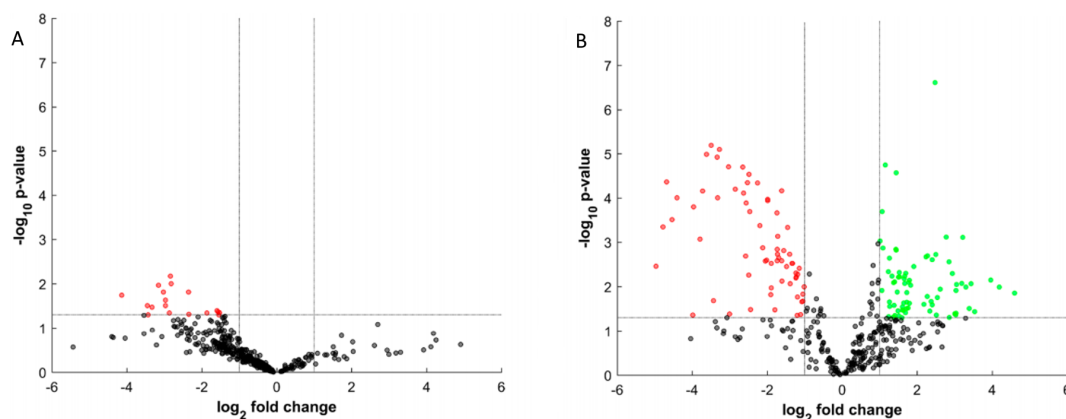
An untargeted metabolomics experiment was performed on archived urine samples from a tea intervention study using LESA-MS. Two  $\mu\text{L}$  of four biological replicates were manually deposited onto the DMA and glass, dried in a fume hood ( $\sim 1$  hour), followed by LESA-MS analysis using MeOH/H<sub>2</sub>O/FA 70:30:0.1 v/v/v as extraction solvent. The acquired spectra were TIC normalised, Pareto scaled and subjected to PCA. The PCA scores plot of urine samples glass shows scattered scores for the samples whilst in the PCA scores plot for sample acquired from DMA are tightly clustered and clearly separated for both groups (Figure 2.4). This revealed that when the spreading

of the sample as well as the extraction cannot be controlled, the technical variance is dominant over the biological variance. Therefore, in this case, no biologically relevant information can be obtained from the acquired data (e.g. biomarkers between two different groups) since the goodness-of-prediction ( $Q^2$ ) is too low. For a relevant biological model, a minimum  $Q^2$  of 0.4 is required [208]. This threshold value is only met when the DMA is used as sample substrate ( $Q^2 = 0.8938$ ).

The control samples analysed on glass are showing, however, low variance between the biological replicates. To some degree, analytes could be extracted from the glass substrate. However, when the extraction solvent is dispensed on glass, there is a high probability that the microjunction is (partially) lost. Therefore, analytes are not or ineffectively extracted from the surface and solvent peaks will dominate the mass spectra. That is what was observed for the control samples on glass. The solvent spreads out on the glass substrate and therefore only the remaining solvent in the tip is analysed. Since solvent spectra are highly similar, the control samples show to be reproducible in the scores, however, they do not bear any biologically-relevant information. On the other hand, the large spread in the *after tea* samples can be explained by the fact that there have been analytes extracted from the glass substrate, however, the solvent recovery was highly variable resulting in a large spread of the scores.



**Figure 2.4:** PCA scores plot for included ions found in urine before (green circles) and after tea intervention (orange squares). (A) LESA-MS on glass surface ( $R^2 = 0.7184$ ;  $Q^2 = -0.1551$ ). (B) LESA-MS on DMA ( $R^2 = 0.9609$ ;  $Q^2 = 0.8938$ )



**Figure 2.5:** Volcano plots for common ions between control and tea intervention urine samples. Green: significantly increasing ion after tea intervention. Red: significantly decreasing ion after tea intervention. (A) Analysis from glass substrate. (B) Analysis from DMA

These findings show potential of using the DMA in combination with LESA-MS as an alternative to conventional methods for *omics* profiling of biological samples. Compared to direct infusion MS, ion suppression caused by a high non-volatile salt content in a biological sample could be reduced due to the dilution effect of the sample in the LESA extraction solvent [209]. Further, shorter run times (*sim* 1 minute for MS profiling) and lower solvent consumption (few microliters) could favour LESA-MS over LC-MS.

## 2.5 CONCLUSION

In this chapter, it is established using simple and cost-effective experimental design that the DMA permits repeatable surface sampling with LESA. As a result, LESA-MS data acquired from biologically different samples is highlighting biological differences between sample groups. The next step is to extend this analysis to proteomics analysis by using the DMA as sample substrate for *in situ* digestion of proteins with subsequent LESA-MS/MS analysis. This will be the focus of the next chapter.

# Optimisation and development of LESA-MS/MS and digestion parameters for *in situ* quantification of proteins

## 3.1 INTRODUCTION

**Q**UANTITATIVE MS-based proteomics has proven to be a valuable tool to quantitatively assess differences in protein profiles between different biological conditions [210]. Classically, MS-based proteomics uses LC-MS/MS for identification and quantification of proteins [76, 132]. However, the sample throughput of LC-MS/MS analysis is low (>1 hour per sample) and would require months to complete the analysis of a polymer microarray (~ 500 samples).

LESA was introduced in Chapter 2 as a suitable technique for automated surface sampling and direct MS analysis. This has permitted sample throughput in the order of minutes which could reduce the analysis time to about 1 day. Combined with a suitable sample substrate (Droplet Microarray [188]; Chapter 2), repeatable LESA sampling has been achieved [211]. The next step is to optimise the experimental and



analysis conditions for proteomics. For reliable protein quantification, it is important that the *in situ* digestion protocol is robust and the LESA and MS parameters permit stable data acquisition, so the technical variation is minimal [212].

In terms of *in situ* digestion, Rao *et al.* [151] proposed an ambient digestion protocol using trypsin and room temperature conditions. Trypsin is the most popular enzyme of choice for MS-based proteomics, however, several other proteases are available [87] and have been unexplored so far for *in situ* digestion. Higher *in situ* digestion efficiencies might be achieved using another enzyme. Each protease cleaves the protein at a specific site and certain cleavage sites might be more accessible *in situ*. Also, a combination of protease might enhance the digestion efficiency due to complementary cleavage [213]. Other adjustable digestion parameters are the temperature and the incubation time. Trypsin has its optimal activity at 37°C [87] which indicates that performing *in situ* digestion under ambient conditions, as done as done by Rao *et al.* [151], might not lead to efficient cleavage of proteins into peptides. So far, optimisation of *in situ* digestion, especially for biomaterial surfaces, is relatively unexplored and requires attention to allow accurate quantification of proteins.

Another point of attention is the MS acquisition strategy. Classically, untargeted proteomics has been performed using DDA [214]. In DDA, peptide ions are selected for fragmentation based on their intensity [100]. However, this stochastic sampling process could underrepresent low-abundant peptide ions. With the development DIA and computational tools it is now possible to perform (un)targeted proteomics without intensity-based of precursor ions [215, 216]. Current reports on the use of LESA-MS/MS for bottom-up proteomics has only focused on DDA, most probably because of its simplicity to implement and analyse with the current (commercial) available software tools. Since it is known that the MS acquisition strategy can affect the reproducibility [214], the use of different LESA-MS/MS acquisition is of interest to investigate and assess the effect on the reproducibility of protein identification and quantification.

In this chapter, parameters for reproducible *in situ* digestion with subsequent LESA-MS/MS analysis are optimised for a robust digestion allowing quantitative assessment

of the amount of protein present on the surface.

### 3.2 AIMS OBJECTIVES

The aim of this thesis chapter is to optimise the *in situ* digestion parameters to permit reproducible protein digestion on the surface. Furthermore, parameters will be optimised for LESA and MS/MS to allow reproducible protein identification and accurate protein quantification.

### 3.3 MATERIALS & METHODS

#### 3.3.1 REAGENTS CHEMICALS

Trypsin (sequencing grade), Trypsin Gold (mass spectrometry grade), trypsin/LysC (mass spectrometry grade), pepsin (sequencing grade), thermolysin, and a MS-compatible yeast extract were acquired from Promega (Southampton, UK). MeOH, ACN and H<sub>2</sub>O (LC-MS grade; CHROMASOLV) was purchased from Riedel-de Haen (Seelze, DE). Dimethyl sulfoxide (DMSO) (Pierce<sup>TM</sup>; LC-MS grade), formic acid (FA) and acetic acid (LC-MS grade; Optima<sup>TM</sup>) were obtained from Fisher Scientific (Loughborough, UK). Amicon 0.5 mL centrifugal filters, recombinant human insulin ( $\geq 98\%$ ), recombinant human *holo*-transferrin ( $\geq 97\%$ ), bovine serum albumin (BSA) ( $\geq 96\%$ ), cytochrome c (Cyt c) (BioUltra;  $\geq 99\%$ ), ammonium bicarbonate (AmBic) (BioUltra;  $\geq 99.5\%$ ), iodoacetamide (IAA) (BioUltra;  $\geq 99.5\%$ ), dithiotreitol (DTT) (BioUltra;  $\geq 99.5\%$ ), and Permax cell culture slides (Nunc®) were obtained from Sigma-Aldrich (Gillingham, UK). Trifluoroacetic acid (TFA) (99.5%) and ammonium hydroxide (ACS reagent) were acquired from Acros Organic (Geel, Belgium). Fibroblast growth factor 2 (FGF-2) ( $>95\%$ ) and transforming growth factor beta 1 (TGF- $\beta 1$ ) ( $>97\%$ ) were obtained from R&D Systems (Abingdon, UK). Droplet Microarrays were obtained from Aquarray GmbH (Karlsruhe, DE)

### 3.3.2 OPTIMISING LESA EXTRACTION PARAMETERS

A 1 mg·mL<sup>-1</sup> BSA, Cyt c, or yeast extract solution was prepared in 100 mM AmBic. Cyt c was introduced to test the method on a smaller, more linear structure type of protein. Yeast proteins were used to test the methodology on a very complex protein mixture. To the solution, 100 mM DTT in 100 mM AmBic was added to an end concentration of 5 mM followed by a 1-hour incubation at 56°C. After the incubation, 100 mM IAA was added to an end concentration of 15 mM and incubated for 30 minutes at room temperature in a dark environment. Next, the solution was transferred to an Amicon 0.5 mL centrifugal unit and washed with 100 mM AmBic for 10 minutes at 13,000 rpm. Thereafter, the purified BSA solution was collected and 5 µL of 0.05 sequencing grade trypsin was added. Digestion was carried out overnight at 37°C followed by storage at -20°C until further use.

#### 3.3.2.1 EXTRACTION SOLVENT VOLUME

The BSA digest was manually pipetted ( $V = 2\ \mu\text{L}$  per replicate) onto a Permanox<sup>TM</sup> cell culture slide and allowed to dry under ambient conditions. The extraction solvent consisted of ACN/H<sub>2</sub>O (1:1 v/v). LESA-MS/MS analysis was performed using the solvent volume parameters shown in Table 3.1. The extraction time was kept constant at 10 seconds without repeated dispensing-aspiration cycles. The variation in the total ion current (TIC) over all MS<sup>1</sup> spectra was calculated and expressed as coefficient of variation (CV).

**Table 3.1:** Solvent volume parameters used for optimisation

Total volume (µL)	Dispensation volume (µL)	Aspiration volume (µL)
1.0	0.8	0.9
1.5	1.3	1.4
2.0	1.8	1.9
2.5	2.3	2.4

### 3.3.2.2 EXTRACTION CYCLES

LESA has the option to repeat the dispensing-aspiration cycle of the extraction solvent. Here, the strategy of Rao *et al.* [151] (7 dispensing/aspiration cycles) was compared to a single dispensing-aspiration cycle to investigate the technical variability. The BSA digest ( $V = 2 \mu\text{L}$  per replicate) was manually pipetted onto a Permanox<sup>TM</sup> cell culture slide and allowed to dry under ambient conditions. Thereafter, the BSA digest was analysed with LESA-MS/MS in triplicate per extraction condition. The variation in TIC between MS<sup>1</sup> scans was used to compare variability between extraction settings.

### 3.3.2.3 EXTRACTION TIME

The same BSA digest ( $V = 2 \mu\text{L}$  per replicate) was manually pipetted onto a clean Permanox<sup>TM</sup> cell culture slide. The digest was allowed to dry under ambient conditions. LESA-MS/MS analysis on the BSA digest was performed using extraction times of 10, 30 and 60 seconds. Each extraction time was tested in triplicate. The most intense BSA peptide ion was used to assess the intensity (peak height) and variability between extraction times.

### 3.3.2.4 EXTRACTION SOLVENT COMPOSITION

The BSA, yeast, or Cyt c digest was pipetted onto a Permanox<sup>TM</sup> cell culture slide and allowed to dry under ambient conditions. LESA-MS/MS analysis was done using the following extraction solvents: ACN/H<sub>2</sub>O/TFA (50:50:0.1 v/v/v), ACN/H<sub>2</sub>O/FA (50:50:0.1 v/v/v), ACN/H<sub>2</sub>O (1:1 v/v), ACN/100 mM ammonium hydroxide (1:1 v/v), 100 mM ammonium hydroxide and ACN/H<sub>2</sub>O/DMSO/FA (50:50:1:0.1 v/v/v/v). Each solvent was tested in triplicate. Intensities (peak height) for peptide matches (MS<sup>1</sup> mass tolerance: 10 ppm) were compared on variation.

### 3.3.3 INFERRING PROTEIN IDENTITIES

The MS-compatible yeast protein extract was reconstituted in 1 mL of 100 mM AmBic, followed by denaturation using 5 mM DTT and 1 hour incubation at 56°C. Thereafter, free thiol groups were alkylated using 15 mM IAA with 30 minutes incubation at room temperature in a dark environment. The solution was purified using a wash with 100 mM AmBic on an Amicon 0.5 mL centrifugal unit. The unit was centrifuged for 10 minutes at 13,000 rpm. 0.05 sequencing grade trypsin was added to the purified solution and incubated overnight at 37°C. After the incubation, the digest was manually pipetted ( $V = 2 \mu\text{L}$  per replicate) in triplicate onto a Permanox<sup>TM</sup> and allowed to dry to the air after which LESA-MS/MS analysis was performed.

Different search algorithms were tested using SearchGUI 3.3.20 [217] and visualisation in PeptideShaker 1.16.45 [218]. SearchGUI supports the following search algorithms: X!Tandem (Vengeance 2015.12.15.2) [104], MyriMatch (2.2.140) [219], MS Amanda [109], MS-GF+ (v2018.04.09) [108], OMSSA [103], Comet [107], Andromeda (1.5.3.4) [105], and Tide [106]. Search algorithms were compared for the time to complete the search, number of peptide spectrum matches (PSMs) and number of unique PSMs. For identification, the false discovery rate (FDR) was set to 1%.

### 3.3.4 LABEL-FREE PROTEIN QUANTIFICATION

The BSA digest generated in Section 3.3.2 was diluted two-fold in 100 mM AmBic. Both 1x and 0.5x solution were two-fold diluted in ACN/0.2% FA. Solutions were analysed via direct infusion MS/MS. To perform label-free quantification, the MaxLFQ algorithm (MaxQuant 1.6.10.43) [220], the most intense tryptic peptide ( $\text{MS}^1$  quantification) and the most intense fragment ion ( $\text{MS}^2$  quantification) were used for comparison.

### 3.3.5 MS ACQUISITION STRATEGIES

Analysis of the BSA digest (Section 3.3.2) was done using DDA (1  $m/z$  isolation window) and DIA (20  $m/z$  isolation window). Data were processed using an in-house developed LESAProteomics MATLAB class ([https://github.com/jorismeurs/LESA\\_Proteomics](https://github.com/jorismeurs/LESA_Proteomics)). Comparison was performed for identification rate (number of identified peptides and identification repeatability), accuracy (repeatability of the measurements), and precision (difference between true and calculated concentration).

### 3.3.6 REPRODUCIBLE *in situ* DIGESTION

To optimise the *in situ* digestion protocol, one of the target proteins for quantification (recombinant human insulin) was used. Insulin stock solution (1 mg·mL<sup>-1</sup>) was prepared in 0.1% v/v acetic acid. Insulin as an excellent protein to study digestion efficiency, since both the single tryptic peptide (GFFYTPK) and the most intense charge state of intact insulin can be captured in a single  $m/z$  range.

Working solutions for sequencing grade trypsin, MS grade trypsin, trypsin/LysC, LysC, Rapid Trypsin, and Rapid Trypsin/LysC were prepared at a concentration of 0.05 in 100 mM AmBic. Protein stock solutions were mixed 1:1 (v/v) and dispensed on a Permax™ cell culture slide and then allowed to dry under ambient conditions. Protease working solutions were tested in triplicate and manually dispensed onto the dry protein spots. Digestion was carried various conditions (time and temperatures, buffers). The ratio between the most intense fully tryptic peptide ion and the intact protein ion were used to assess the digestion efficiency [221].

### 3.3.7 QUANTIFICATION OF TARGET PROTEINS

Individual 2-fold dilution series were made from stock solutions of insulin (1 mg·mL<sup>-1</sup> in 1% acetic acid), transferrin (1 mg·mL<sup>-1</sup> in 100 mM AmBic), FGF-2 (250 µg·mL<sup>-1</sup> in 100 mM AmBic), and TGF-β1 (50 µg·mL<sup>-1</sup> in 100 mM AmBic). These proteins chosen here represent the target proteins for studying adsorption on synthetic polymer substrates.

Dilutions were introduced onto a Droplet Microarray via the rolling droplet technique (Figure 3.1) for equal distribution of the solutions [188]. The rolling droplet technique starts by forming a large droplet on the end of a glass Pasteur pipette. This droplet is then dragged across the DMA. Due to extreme difference in wettability, individual droplets are formed on the surface. The difference in droplet volume was found to be smaller than 5% [188]. After drying under ambient conditions, *in situ* digestion and LESA-MS/MS were carried out under optimised conditions.



**Figure 3.1:** Visualisation of the rolling droplet technique as described by Popova *et al.* [188]. The extreme difference in wettability of the pattern allows formation of individual droplets on the array

### 3.3.8 ANALYSIS

All optimisation work was carried using a TriVersa Nanomate (Advion Biosciences, Ithaca, NY) coupled to a Q Exactive Orbitrap mass spectrometer (Thermo Scientific, San Jose, CA). Thermo .RAW files were converted to .mzXML format (32-bit) using ProteoWizard [199]. Statistical analysis was carried out in MATLAB R2017a (The MathWorks, Inc., Natick, MA) on a Lenovo Thinkpad x260 with Microsoft Windows 10 Enterprise as operating system. For sequence database searching, the precursor and fragment ion mass tolerance were set to 10 ppm and 0.02 Da respectively.

## 3.4 RESULTS & DISCUSSION

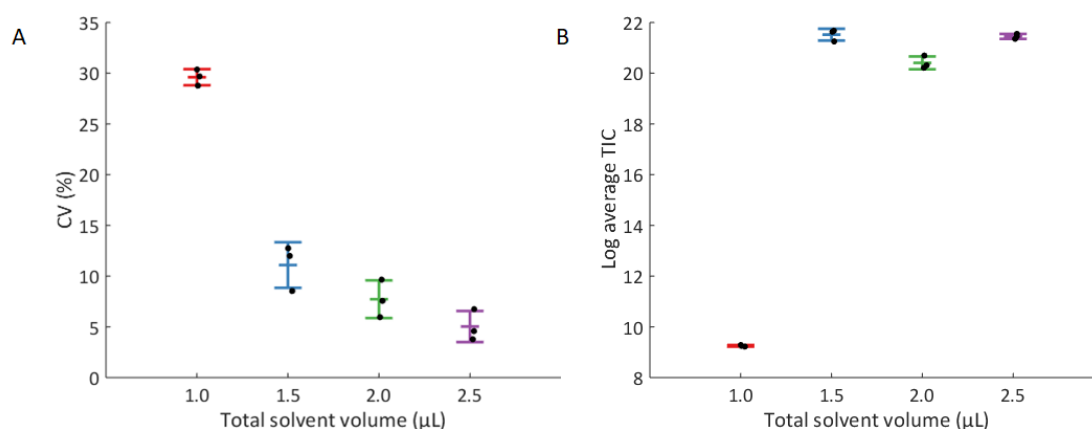
### 3.4.1 OPTIMISING LESA & MS PARAMETERS

For optimisation of the LESA extraction parameters, BSA was used because of its size (66 kDa), so many peptides can be generated from the intact protein, its costs (relatively

cheap compared to other analytical grade protein standards) and its wide use for optimisation so it will allow for comparison with previous studies. For more in-depth assessment of the extraction parameters, a commercial yeast digest (Promega, Southampton, UK) was used. To discover which MS acquisition strategy is most suitable for identification, the target proteins for array screening experiments (insulin, transferrin, FGF-2 and TGF- $\beta$ 1 ) were used. Results of subsequent optimisation experiment are given in the following sections.

### 3.4.1.1 EXTRACTION SOLVENT VOLUME

Different volumes of ACN/H<sub>2</sub>O 1:1 v/v were tested in triplicate on a BSA digest manually deposited onto a Permanox<sup>TM</sup> cell culture slide. Permanox<sup>TM</sup> slides are hydrophobic and also permit sufficient control of the extraction solvent. Since the DMAs are expensive products, optimisation experiments were performed on the cheaper alternative Permanox<sup>TM</sup>. Using 1.0  $\mu$ L was found to be a too small volume, i.e. a low and unstable signal was obtained during infusion. The signal stability increased significantly by increasing the solvent volume to 2.5  $\mu$ L (Kruskal-Wallis:  $p = 0.0232$ ; Figure 3.2A). The intensity increased significantly after increasing the solvent to 1.5  $\mu$ L ( $p = 1.041 \times 10^{-12}$ ) and remained at the same level after increasing to 2.0  $\mu$ L and 2.5  $\mu$ L (Figure 3.2B).



**Figure 3.2:** Optimising solvent volume for LESA-MS/MS analysis of tryptic peptides. (A) TIC deviation expressed as CV. (B) Average TIC per solvent volume

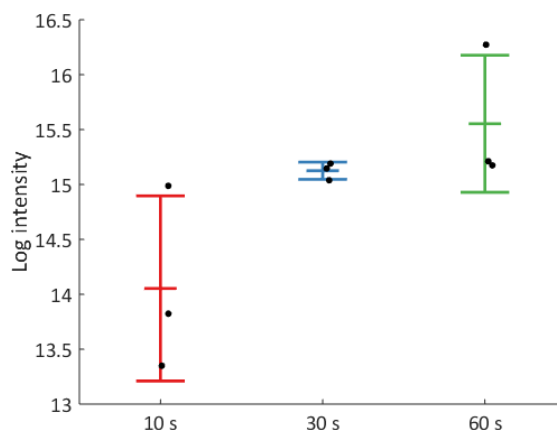


### 3.4.1.2 EXTRACTION CYCLES

It was found that multiple dispensing aspiration cycles decrease signal stability and therefore increasing the technical variance. In order to recover the solvent from the surface during the aspiration, the volume for aspiration ( $V_{asp}$ ) is in general greater than the dispensed volume ( $V_{disp}$ ). Though the mixing of the analytes will be improved during multiple cycles, every cycle solvent is lost on the surface and since  $V_{asp}$  is greater than  $V_{disp}$ . In addition, an increasing air gap will be formed at the end of the tip. As a consequence, no signal is obtained during the first part of the acquisition since the solvent has not reached the end of the tip yet. Therefore, no MS data could be acquired to compare signal stability between single and multiple extraction cycles. However, the empirical observation that no data could be generated using multiple extraction cycles led to the decision to use only a single dispensing-aspiration cycle.

### 3.4.1.3 EXTRACTION TIME

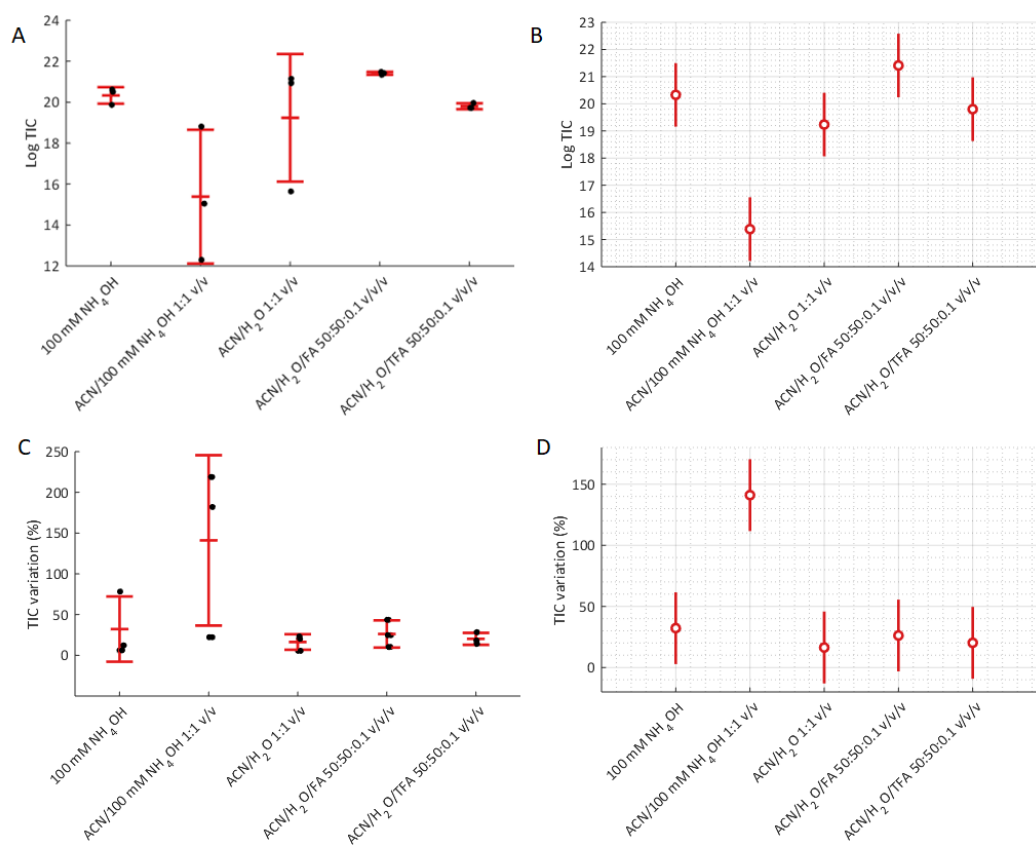
When investigating the influence of the extraction time on the peptide intensity, a higher intensity for QTALVELLK (2+;  $m/z$  507.8138) was observed when increasing the extraction time (Figure 3.3), though, the difference was not found to be statistically significant (ANOVA:  $p = 0.056$ ). When investigating the reproducibility, it was found that extraction for 30 seconds led to the lowest variability in peptide intensity. However, the difference in standard deviation of the peptide intensity between extraction times was not found to be statistically significant (Levene's test:  $p = 0.0646$ ). From these findings can be derived that similar extraction efficiencies can be obtained without comprising the reproducibility regardless of the extraction time



**Figure 3.3:** The impact of extraction time on the intensity of the tryptic BSA peptide QTALVELLK (2+;  $m/z$  507.8138)

#### 3.4.1.4 EXTRACTION SOLVENT COMPOSITION

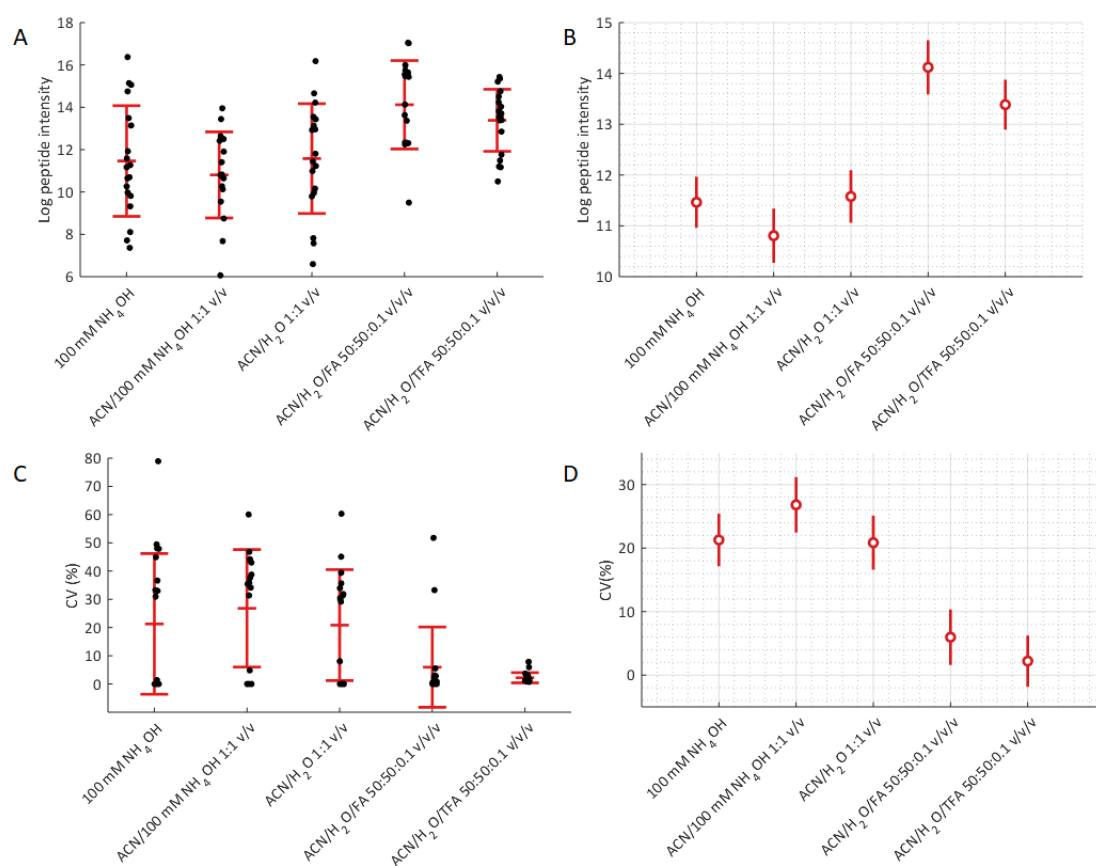
A BSA digest on a Permanox™ cell culture slide was analysed using LESA-MS/MS with a number of different extraction solvents. First, the signal stability (TIC variation) was investigated. A stable signal is important for reliable quantification. All the tested extraction solvents were found to perform equal in terms TIC intensity, except for ACN/100 mM  $\text{NH}_4\text{OH}$  1:1 (v/v), which signal was about 2 orders of magnitude lower (ANOVA:  $p = 0.0383$ ; Figure 3.4A-B). ANOVA on the TIC variation showed equal performance of each extraction solvent ( $p = 0.0603$ ), though, through *post hoc* analysis the variation in TIC intensity was found to be greater for ACN/100 mM  $\text{NH}_4\text{OH}$  1:1 v/v (Figure 3.4C-D).



**Figure 3.4:** Comparison of the signal stability for different extraction solvents. (A) Jitter plot for the average TIC per extraction solvent. (B) Statistical analysis (ANOVA + Tukey-Kramer *post hoc* analysis) for comparison of the TIC intensity. (C) TIC variation for different extraction solvents. (D) Statistical analysis (ANOVA + Tukey-Kramer *post hoc* analysis) for comparison of the TIC variation.

Besides a stable TIC, the signal stability for peptides is also an important parameter to assess for accurate quantification. Overall, the peptide signal was more intense when using FA or TFA as additive (ANOVA:  $p = 4.19 \times 10^{-5}$ ; Figure 3.5A-B). For reliable quantification of proteins, the CV of the peptide signal should be below 20% [132]. All CVs were below 20% when TFA was used as additive and for most of the peptides when using FA as an additive. Both solvents (ACN/H<sub>2</sub>O /FA 50:50:0.1 (v/v/v) and ACN/H<sub>2</sub>O /TFA 50:50:0.1 (v/v/v)) had a better peptide signal reproducibility than other tested solvents (ANOVA:  $p = 0.0001$ ; Figure 3.5C-D). In terms of sensitivity and signal stability, both extraction solvents outperform ACN/H<sub>2</sub>O 1:1 (v/v) which was previously proposed as optimal solvent to sample from biomaterial surfaces [151]. The use of

ammonium hydroxide was explored to assess whether an alkaline solution would be more suitable (i.e. signal stability). It has been reported that alkaline solution could improve the solubility of more hydrophobic components and cleaner (less solvent adduct ions) mass spectra can be obtained due to reduced matrix interference [222]. The hydrophobicity of peptides can be expressed as grand average of hydropathy (GRAVY) score. The higher the GRAVY score, the more hydrophobic the peptide is [223]. Based on current data, an ammonium hydroxide solution does not improve the signal intensity for hydrophobic peptides compared to ACN/H<sub>2</sub>O 1:1 (v/v) (Table 3.2).



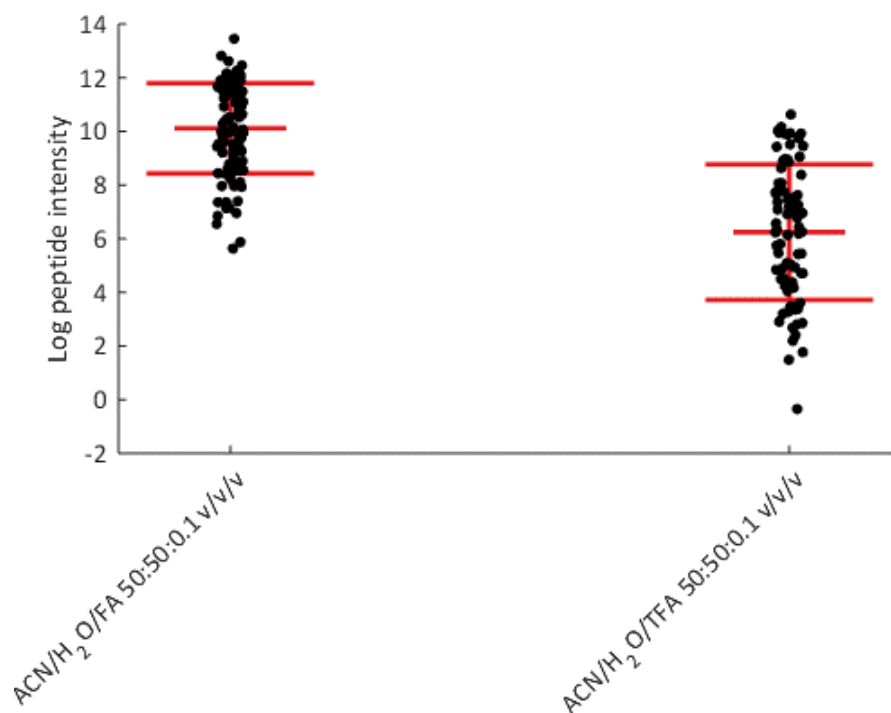
**Figure 3.5:** Comparison of BSA peptide signal intensity for different extraction solvents. (A) Jitter plot for peptide signal intensity distribution. (B) ANOVA followed Tukey-Kramer *post hoc* analysis for comparison of peptide signals. Non-overlapping intervals are significantly different. (C) Jitter plot for variation in peptide signal intensity. (D) ANOVA followed Tukey-Kramer *post hoc* analysis for comparison of variation in peptide signal. Non-overlapping intervals are significantly different

**Table 3.2:** Signal intensity for hydrophobic BSA peptides when analysed with 100 mM NH<sub>4</sub>OH compared to ACN/H<sub>2</sub>O 1:1 (v/v)

Peptide	GRAVY score	Log intensity	Fold change
QTALVELLK	0.64	13.49	1.03
HPEYAVSVLLR	0.26	14.75	0.96
AEFVEVTK	0.18	14.52	1.02
LVNELTEFAK	0.13	13.14	0.98
LGEYGFQNALIVR	0.29	14.95	1.11
KQTALVELLK	0.19	14.79	0.96

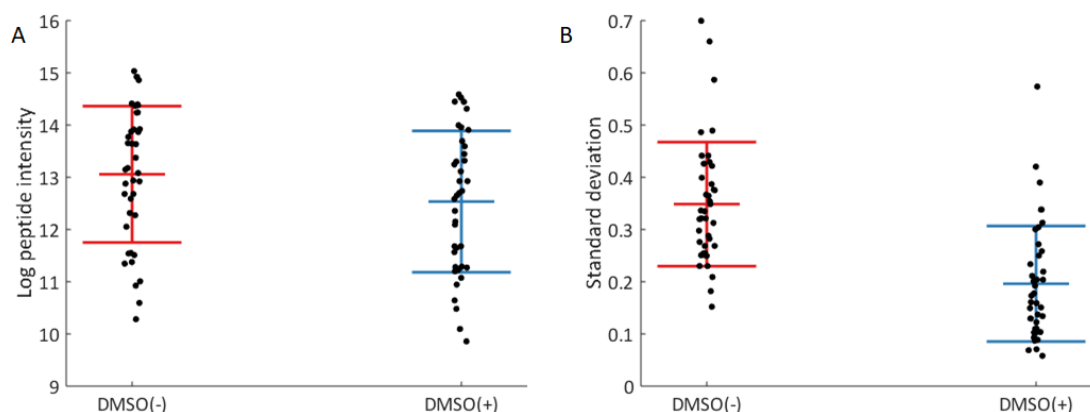
Fold changes were calculated as ratio between the mean peptide signal intensity using ACN/H<sub>2</sub>O 1:1 (v/v) and the mean peptide signal intensity using 100 mM NH<sub>4</sub>OH as extraction solvent

In order to determine which additive is more suitable (least variation; high identification rate), a yeast digest on a Permax™ cell culture was analysed in triplicate. TFA is normally used as an ion-pairing agent for improved chromatographic separation of peptides [224]. However, TFA is also well-known for its suppressing effect for the ionisation of peptides [225]. This became evident after the sequence database search for all LESA-MS/MS data. The difference in the number of identified peptides (FA: 107; TFA: 81) demonstrated the potential consequences of using TFA as an additive. Further investigation of the peptide intensities revealed a significant reduced peptide ion intensity when using TFA as an additive (Student's *t*-test :  $p < 0.001$ ; Figure 3.6). Therefore, the use of TFA as an electrospray additive should be omitted for LESA-MS/MS analysis of peptides.



**Figure 3.6:** Jitter plot for comparison of the yeast peptide intensities for FA and TFA as extraction solvent additive

In order to improve the solubility of hydrophobic peptides, DMSO was introduced as alternative for TFA. DMSO is known for improving the solubility of hydrophobic compounds and has been reported to improve the electrospray signal of peptides [226–229]. A Cyt c digest was analysed from Permanox<sup>TM</sup> cell culture slides. Cyt c was used here since it is easier to digest and would therefore allow proper comparison of peptide intensities in this experiment. DMSO was added to 1% (v/v) in the extraction solvent. No difference in the overall peptide intensity distribution (Student's *t*-test :  $p = 0.0827$ ; Figure 3.7A) was observed, indicating no enhanced peptide intensity. However, it was found that the use of 1% (v/v) DMSO in ACN/H<sub>2</sub>O /FA 50:50:0.1 v/v/v improved the signal repeatability (Levene's test:  $p < 0.001$ ; Figure 3.7B) compared to ACN/H<sub>2</sub>O /FA 50:50:0.1 v/v/v.



**Figure 3.7:** Investigating the benefit of DMSO as additive for the LESA extraction solvent. (A) Peptide intensity distribution with and without DMSO. No difference in signal intensity was observed with or without 1% (v/v) DMSO (Student's *t*-test :  $p = 0.0827$ ). (B) Peptide intensity variation with and without DMSO. Signal repeatability significantly improved using 1% (v/v) DMSO as an additive (Levene's test:  $p < 0.001$ )

### 3.4.2 SEARCH ALGORITHM SELECTION FOR MAXIMUM PEPTIDE IDENTIFICATION

The yeast protein digest was manually pipetted onto a Permax™ cell culture slide and analysed in triplicate with LESA-MS/MS. Thermo .RAW files were submitted to SearchGUI [217] followed by visualisation in PeptideShaker [218]. The yeast digest was used to be able to assess the performance of the search algorithms on a complex proteome. The results from each search algorithm are shown in Table 3.3. It is well-known that the use of different search algorithms leads to different results [111]. Andromeda, Comet and MS-GF+ required significantly more time to complete the search compared to the other search algorithms (ANOVA:  $p < 0.001$ ). Though, the benefit of MS-GF+ is the identification of peptide-spectrum matches (PSMs) which could not be detected by any other search algorithm. A PSM is defined as a MS/MS spectrum that is matched to a peptide. Unique PSMs (PSMs not identified with any other algorithm) were also obtained using OMSSA, X!Tandem, MyriMatch, and Tide. The gain of unique PSMs with MS-GF+, MyriMatch and Tide is only minimal, but MS-GF+ allows the identification of more PSMs (ANOVA:  $p < 0.001$ ). Based on the results of this experiment,

X!Tandem, OMSSA and MS-GF+ were selected as search algorithms for the analysis of LESA-MS/MS data.

**Table 3.3:** Performance of search algorithms for identifying peptide from a yeast digest analysed by LESA-MS/MS

Search algorithm	Search time (s)	Total PSMs	Unique PSMs
OMSSA	13 ± 1	697 ± 25	25 ± 4
X!Tandem	9 ± 0	658 ± 25	19 ± 2
Andromeda	115 ± 12	621 ± 24	0 ± 0
MS Amanda	16 ± 1	633 ± 24	0 ± 0
MS-GF+	79 ± 4	736 ± 37	3 ± 1
Comet	155 ± 11	732 ± 40	0 ± 0
MyriMatch	16 ± 1	724 ± 38	1 ± 1
Tide	27 ± 1	627 ± 26	1 ± 1

The search time refers as the time for the algorithm to complete the search and identification as logged by SearchGUI

### 3.4.3 REPRODUCIBLE AND ACCURATE PROTEIN IDENTIFICATION & QUANTIFICATION

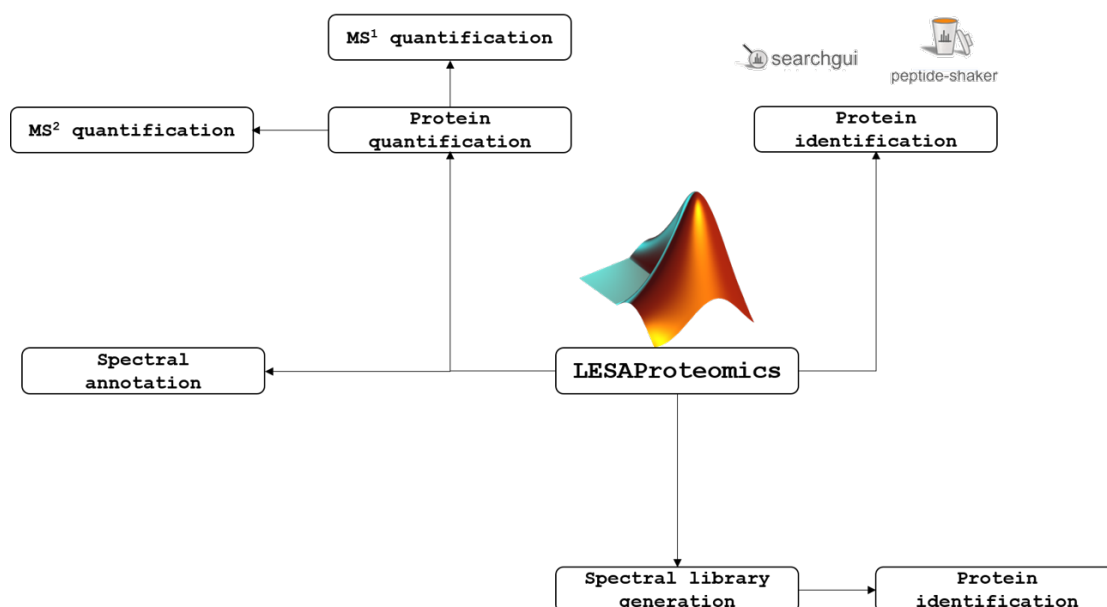
#### 3.4.3.1 A CUSTOM MATLAB TOOL FOR QUANTIFYING PROTEINS FROM LESA-MS/MS AND DIRECT INFUSION MS/MS DATA

Two dilutions of a BSA digest were analysed via direct infusion MS/MS for quantitative analysis. To use the MaxLFQ algorithm, data needs to be analysed in MaxQuant [220]. MaxQuant is a popular open-source processing platform for proteomics data acquired on a LC-MS/MS system. The platform has been used previously for quantification of direct infusion MS/MS data acquired using the TriVersa Nanomate [230]. In addition to the long processing times ( $\pm 30$  minutes for 6 files (200 MB/file)), it was often observed that MaxQuant crashes without any error report. Furthermore, the results obtained via MaxQuant's MaxLFQ algorithm do not reflect the concentration difference (fold change) between both samples. The MaxLFQ algorithm gives  $3.46 \times 10^9$  LFQ



intensity for the 1x solution and 0 LFQ intensity for the 0.5x solution. Additional quantitative output in MaxQuant is the summed peptide intensity for all MS/MS identified peptides. Using the summed intensity of all identified BSA peptides, the calculated fold change is 124 whilst this should be 2.

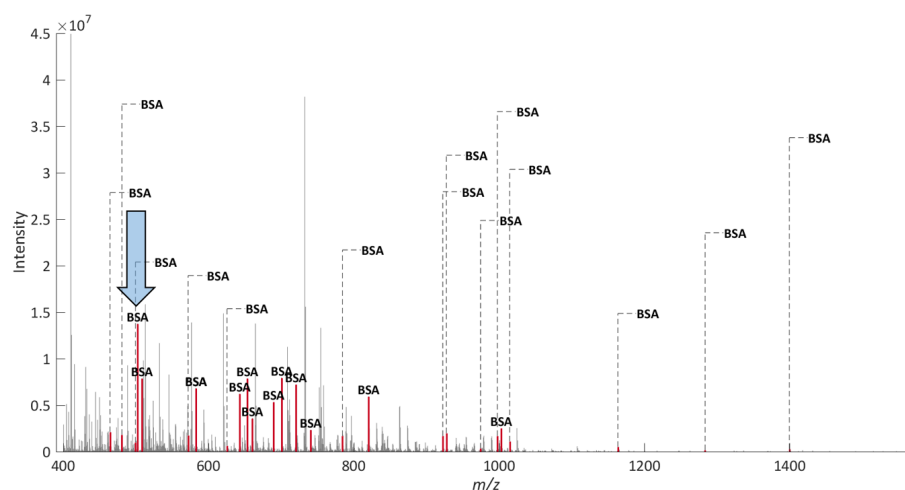
Since (open source) software are specifically designed for LC-MS/MS data, processing LESA-MS/MS or direct infusion MS/MS data quantitatively with existing tools would be rather problematic, due to the absence of peptide elution profiles (peaks). For that reason, a MATLAB class was constructed for identification and quantification of proteins. LESAProteomics (Figure 3.8) uses the command line version of SearchGUI [217] and PeptideShaker [218] for protein identification. Furthermore, the MATLAB class permits annotation of MS<sup>1</sup> and MS<sup>2</sup> spectra. In addition, quantification can be performed at both MS<sup>1</sup> and MS<sup>2</sup>. The class has been made openly available via [https://github.com/jorismeurs/LESA\\_Proteomics](https://github.com/jorismeurs/LESA_Proteomics).



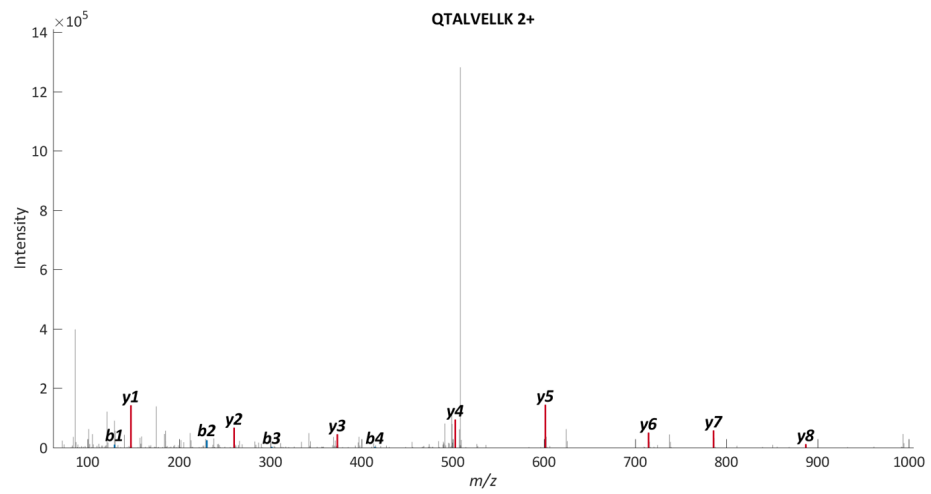
**Figure 3.8:** The LESAProteomics MATLAB class for processing of proteomics data acquired via LESA-MS/MS

Quantification of BSA was performed based on the most intense peptide ion and the most intense fragment ion. BSA peptides were identified through SearchGUI and

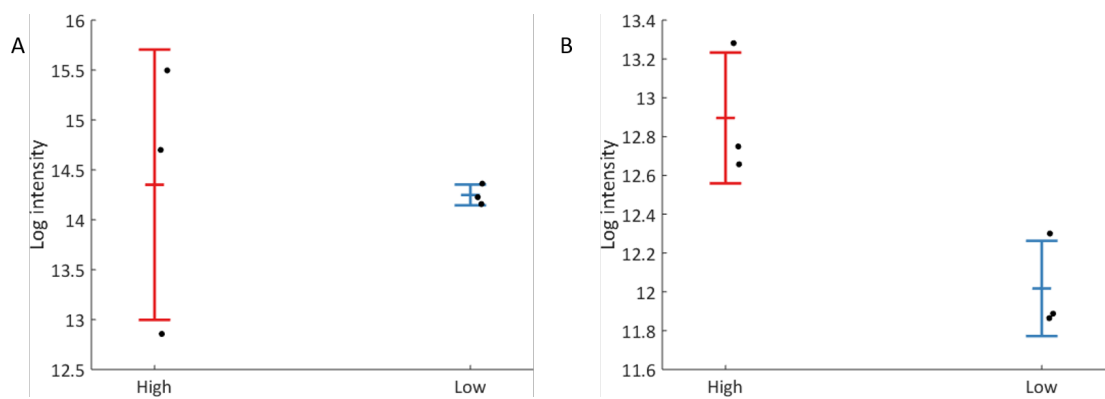
PeptideShaker. It was found that QTALVELLK (2+;  $m/z$  507.8142) was the most intense peptide ion (Figure 3.9). For the same peptide, the most intense fragment ion ( $y_5$ ) was extracted (Figure 3.10). Peptide intensities were log transformed to account for technical variability [231]. Analytical requirements for quantification of proteins is a CV below 20% [216]. At  $MS^1$  level, the CV for respectively BSA 1x and BSA 0.5x was 9.4% and 0.7% whilst at  $MS^2$  the CVs were 2.6% and 2.0% respectively (Figure 3.11). Both quantification strategies are well within the limits for accurate quantification. No significant difference in variation was found between both strategies for BSA 1x ( $F$ -test:  $p = 0.1166$ ) and BSA 0.5x ( $F$ -test:  $p = 0.3054$ ). The calculated fold change was found to be between for  $MS^1$  (1.10) and  $MS^2$  (2.41) quantification. Based on this calculation,  $MS^2$  quantification produces more accurate results and therefore reflect the actual concentration different better compared to  $MS^1$ . To confirm this observation, an additional experiment was run to assess whether  $MS^1$  or  $MS^2$  using two of the target proteins for the future array screening experiments (insulin and transferrin) after further optimisation of the digestion parameter (Section 3.4.5).



**Figure 3.9:** Annotated  $MS^1$  spectrum for BSA peptides. QTALVELLK is indicated by the black arrow



**Figure 3.10:** Annotated MS<sup>2</sup> spectrum for the most intense BSA peptide (QTALVELLK 2+)



**Figure 3.11:** Jitter plot for comparing (A) MS<sup>1</sup> and (B) MS<sup>2</sup> quantification

**Table 3.4:** Comparison of MS protein quantification strategies

Parameter	MS <sup>1</sup>	MS <sup>2</sup>
CV 1x	9.4%	2.6%
CV 0.5x	0.7%	2.0%
Fold change	1.10	2.41

### 3.4.3.2 SELECTING THE MS ACQUISITION STRATEGY

The BSA digest was analysed via direct infusion MS/MS. The Orbitrap mass spectrometer was operated in DDA and DIA mode to investigate which acquisition mode results in the most reproducible peptide identification and peptide signal intensity.

With classic DDA analysis of peptides, one has to deal with the stochastic nature of precursor selection for fragmentation [214]. Therefore, MS/MS spectra of peptides might not be acquired especially in the case of low abundance. In addition, peptide signals can be severely suppressed due to increased ion suppression as a result of the absence of chromatography [232]. As an alternative, DIA can be employed which overcomes the biased precursor sampling. When using DIA, all target proteins (insulin, transferrin, FGF-2 and TGF- $\beta$ 1 ) were identified using Skyline [117], whilst operating the MS in DDA mode, the detection rate was reduced (Table 3.5). In order to be able to perform quantification, identification of the peptides of interest should be robust. The other advantage of operating in DIA is the reduced analysis time per sample ( $\sim$  1 minute per sample).

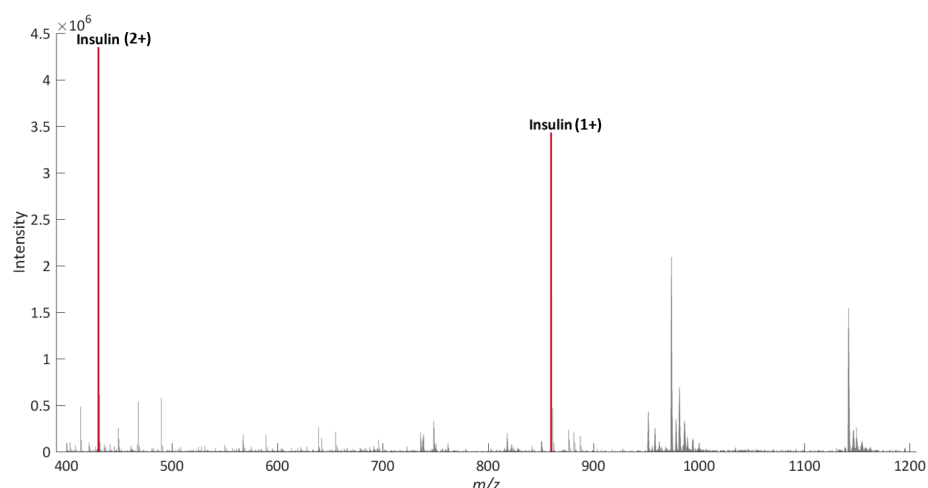
**Table 3.5:** Identification rate of target proteins using DDA and DIA

Protein	DDA	DIA
Insulin	0/3	3/3
Transferrin	3/3	3/3
FGF-2	1/3	3/3
TGF- $\beta$ 1	1/3	3/3

### 3.4.4 ROBUST *in situ* DIGESTION

So far, all parameters have been optimised using model proteins. In order to assess the performance of the protein digestion, one of the target proteins (insulin) was used. The advantage of insulin is that its tryptic peptides and intact protein charge states can be captured in a single, MS-compatible  $m/z$  range [233], which makes it ideal to monitor the peptide ions as well as the intact protein ions.

Several enzymes were selected for assessment of the *in situ* digestion efficiency of insulin. First, the most intense tryptic peptides was derived from a standard *in situ* digests. Since recombinant human insulin has only one fully tryptic peptide (GFFYTPK 2+;  $m/z$  430.2212; Figure 3.12), no further investigation is necessary.

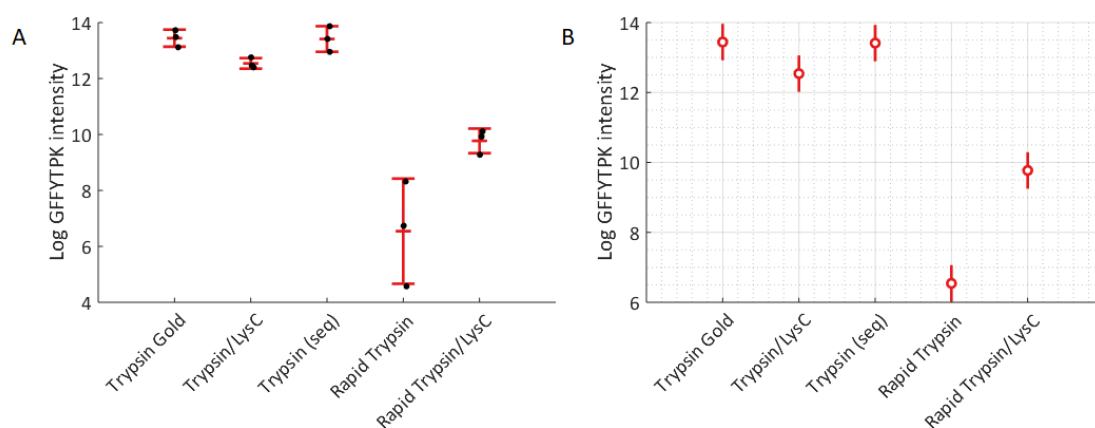


**Figure 3.12:** Annotated MS<sup>1</sup> spectrum for tryptic peptides generated from recombinant human insulin. The most intense peak ( $m/z$  430.2212) represents GFFYTPK (2+) whilst the other annotation is the singly charged peptide ion for GFFYTPK (1+;  $m/z$  859.4342)

Initially, the *in situ* digestion conditions listed in Table 3.6 were used based on previous research [151] and manufacturer recommendations. Peptide ion intensities were extracted from the spectra using MATLAB. GFFYTPK (1+: 859.4342; 2+:  $m/z$  430.2212) was readily detected, however, the intact protein ion was not found for all replicates. In order to statistically assess the digestion reproducibility, the peptide intensity of GFFYTPK (2+) was used as alternative. Higher peptide intensities were observed when using sequencing grade trypsin, Trypsin Gold (MS grade) or the Trypsin/LysC compared to the heat-stable Rapid Trypsin and Rapid Trypsin/LysC mix (ANOVA:  $p = 1.02 \times 10^{-5}$ ; Figure 3.13). Furthermore, digestion using Rapid Trypsin was found to be less reproducible based on the variation in signal intensity for GFFYTPK (Bartlett's test:  $p = 0.0263$ ).

**Table 3.6:** Digestion conditions for used proteases

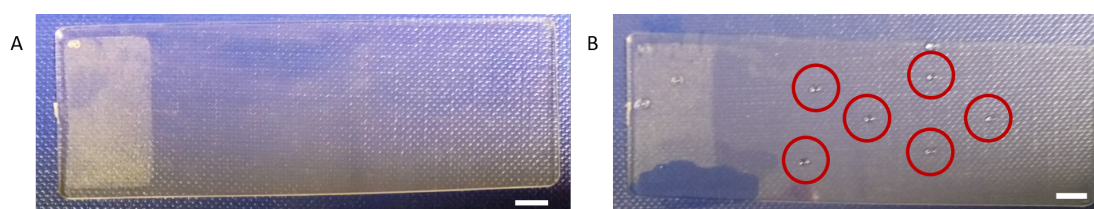
Protease	Temperature (°C )	Time (h)
Sequencing grade trypsin	37	18
Trypsin Gold (MS grade)	37	18
Trypsin/LysC	37	18
Rapid Trypsin	70	3
Rapid Trypsin/LysC	70	3



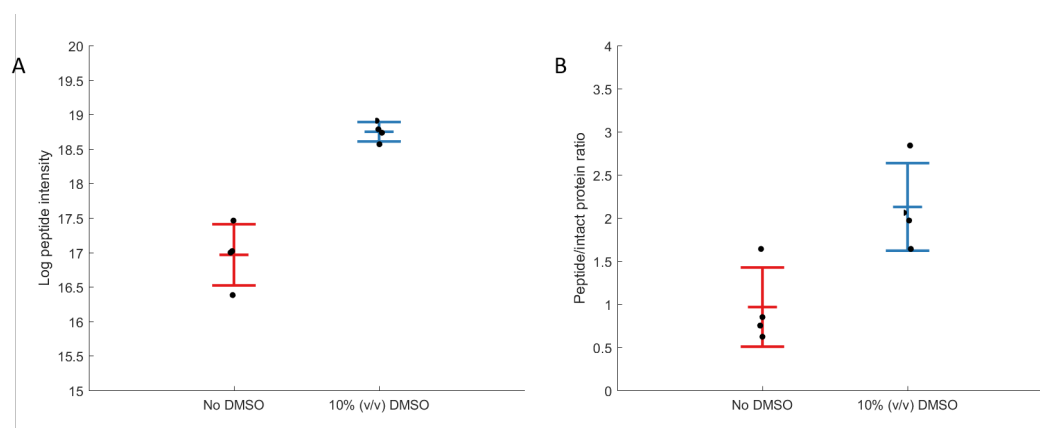
**Figure 3.13:** Comparing *in situ* digestion performance for recombinant human insulin with different proteases. (A) Jitter plot for the peptide intensity distribution of GFFYTPK. (B) ANOVA with subsequent *post hoc* analysis (Tukey-Kramer test) for statistical comparison of peptide intensities. Non-overlapping intervals are significantly different

The reason for variability in detection and peptide intensity could be due to the evaporation of the digestion buffer during analysis. It was noted that all spots for digestion were dried after incubation, despite performing the digestion in a humid environment (Figure 3.14A). Adding a wet tissue to create a humid environment [151] did not prevent evaporation. In order to prevent evaporation of the digestion buffer, DMSO (10% v/v) was added to the digestion buffer. At macroscopic level, the digestion buffer containing 10% (v/v) DMSO did not seem to evaporate during incubation for 2 hours at 37°C (Figure 3.14B). When DMSO was added to the digestion buffer to a final concentration of 10% (v/v), it was found that both the peptide ion intensity

for GFFYTPK (Figure 3.15A; Student's  $t$ -test :  $p = 2.58 \times 10^{-4}$ ) and the peptide/intact protein ratio significantly increased (Figure 3.15B; Student's  $t$ -test :  $p = 0.0146$ ). The digestion reproducibility was found to be similar between the two conditions (Levene's test:  $p = 0.9251$ ). DMSO has been reported to improve solubilisation of more hydrophobic proteins [234] and as an enhancer for trypsin activity [235], explaining a more efficient digestion when using a digestion buffer containing DMSO and 100 mM AmBic (1:9 v/v).



**Figure 3.14:** The addition of DMSO prevented the evaporation of the digestion buffer (B) compared to the standard 100 mM AmBic digestion buffer. After 2 hours incubation at 37°C, the deposited 2 µL is still present in B (indicated by the red circles) whilst in A no solution is visible. This indicates the addition of DMSO is required to prevent the digestion to stop before completion Scale bar: 5 mm



**Figure 3.15:** Investigating the effect of DMSO as an additive to prevent evaporation of the digestion buffer. (A) Jitter plot for comparison of the GFFYTPK peptide intensity. (B) Jitter plot for comparison of the digestion efficiency measured by the peptide/intact protein ratio

The effect of DMSO on the digestion was further investigated. A selection of the previous proteases were used for *in situ* digestion under slightly different conditions

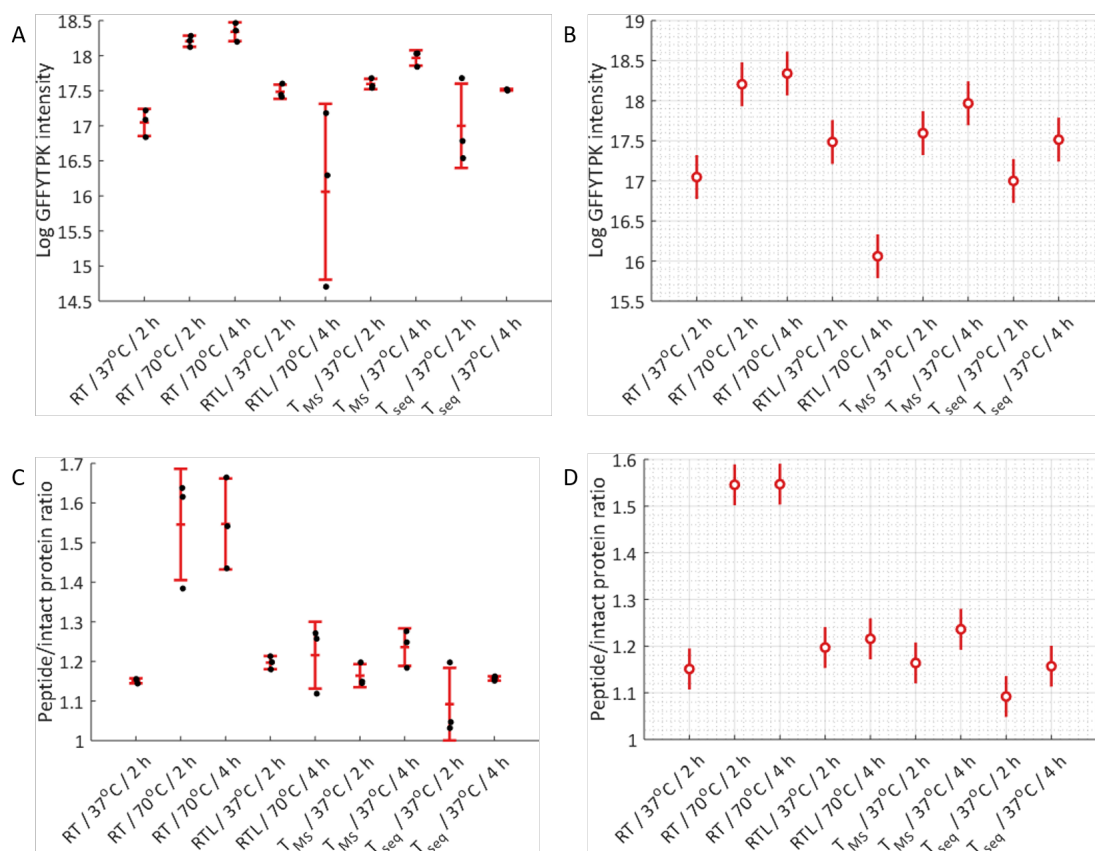
which were reported to be more optimal for LESA-MS/MS analysis of tryptic peptides [236] (Table 3.7).

**Table 3.7:** Digestion conditions used for the second round of experiments using DM-SO/100 mM AmBic 1:9 v/v

Protease	Temperature (°C )	Time (h)
Sequencing grade trypsin	37	2
Sequencing grade trypsin	37	4
Trypsin Gold (MS grade)	37	2
Trypsin Gold (MS grade)	37	4
Rapid Trypsin	37	2
Rapid Trypsin	70	2
Rapid Trypsin	70	4

Higher peptides intensities (ANOVA:  $p = 0.0005$ ; Figure 3.16A-B) as well as digestion efficiency is achieved using Rapid Trypsin at 70°C (ANOVA:  $p = 0.0003$ ; Figure 3.16C-D). The incubation time was not of influence. The use of all other proteases resulted in similar digestion efficiency. Despite the improved digestion using Rapid Trypsin, its variability between replicates was found significantly larger compared to the other conditions (Levene's test:  $p = 0.00481$ ). In order to perform reliable quantitative protein analysis, the variation caused by technical handling (e.g. digestion) should be kept as low as possible. Therefore, low deviation in digestion efficiency should weigh higher in selecting the protease and conditions than the actual digestion efficiency. All the digestion protocols performed similar in terms of repeatability (Levene's test:  $p = 0.1778$ ). As a results, the selection of the digestion conditions for further studies was based on the lowest variation in digestion efficiency which was observed for *in situ* digestion with sequencing grade trypsin and incubation for 4 hours at 37°C (Table 3.8).





**Figure 3.16:** Comparison of the digestion efficiency for insulin under different digestion conditions. (A) Intensity (peak height) distribution for GFFYTPK 2+ for each condition. (B) ANOVA followed by Tukey's *post hoc* correction for statistical comparison of the peptide intensities. Non-overlapping intervals are significantly different. (C) Distribution of the digestion efficiency for insulin (peptide-intact protein ratio) for each digestion condition. (D) ANOVA followed by Tukey's *post hoc* correction for statistical comparison of the digestion efficiency. Non-overlapping intervals are significantly different.

**Table 3.8:** Variation in digestion efficiency per protein for different digestion conditions

Protease	Temperature (°C )	Time (h)	Trf (CV%)	Ins (CV%)	FGF-2 (CV%)	TGF- $\beta$ 1 (CV%)
Sequencing grade trypsin	37	2	9.2	17.0	6.5	16.8
Sequencing grade trypsin	37	4	2.7	1.27	9.3	9.5
Trypsin Gold (MS grade)	37	2	22.0	3.3	23.7	22.5
Trypsin Gold (MS grade)	37	4	6.2	17.0	6.5	16.8
Rapid Trypsin	37	2	32.6	16.8	19.3	34.8
Rapid Trypsin	70	2	7.26	16.86	19.80	46.41
Rapid Trypsin	70	4	14.4	20.2	19.8	16.2

### 3.4.5 PROTEIN QUANTIFICATION WITH OPTIMISED PARAMETERS

With the digestion parameters optimised, serial diluted insulin and transferrin solutions were prepared for quantitative studies. Insulin and transferrin are two of the target proteins for this thesis and were therefore used for further optimisation. All dilutions were manually deposited in triplicate on a Droplet Microarray ( $\varnothing$ 1.414 mm). The solvent volumes for LESA were slightly altered. Empirical observation showed that improved solvent recovery from the surface was achieved when the total volume was set to 3.0  $\mu$ L and the dispensation and aspiration volume were set to 1.5  $\mu$ L and 2.0  $\mu$ L, respectively. Reproducible solvent recovery is required to reduce technical variance in the data.

For both insulin and transferrin, the most intense MS/MS identified peptide was selected. Since from insulin only one tryptic peptide can be generated (GFFYTPK), this peptide was used for quantification. Two charge states were observed for GFFYPTK,

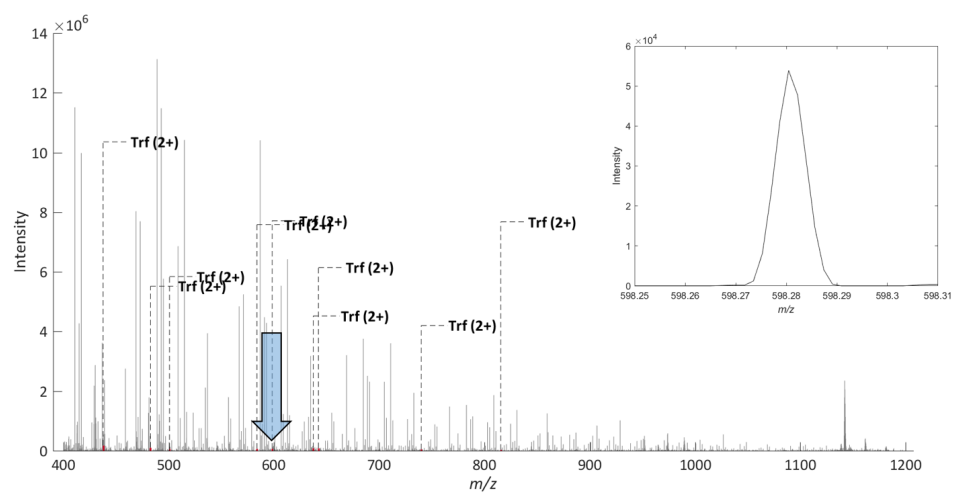
from which the doubly charged peptide ion was found to be the most intense ( $m/z$  430.2210; Figure 3.12). For transferrin, DSGFQMNQLR (2+;  $m/z$  598.2805) was found to be the most intense peptide ion (Figure 3.17). The fragmentation spectra for GFFYPTK (2+) and DSGFQMNQLR (2+) are shown in Figure 3.18. Both MS<sup>1</sup> and MS<sup>2</sup> quantification strategies were compared for both proteins. Proper choice of the peptide for quantification is important since not all peptides bear quantitative properties, i.e. not all peptide ions show a concentration-dependent change in intensity [237]. The quantification was assessed based on relative variation between replicates (CV) and the accuracy (% error) calculated via Equation 3.1 in which  $FC_{obs}$  and  $FC_{true}$  are the measured and true fold change, respectively.

$$\%error = \frac{|FC_{true} - FC_{obs}|}{FC_{true}} \times 100\% \quad (3.1)$$

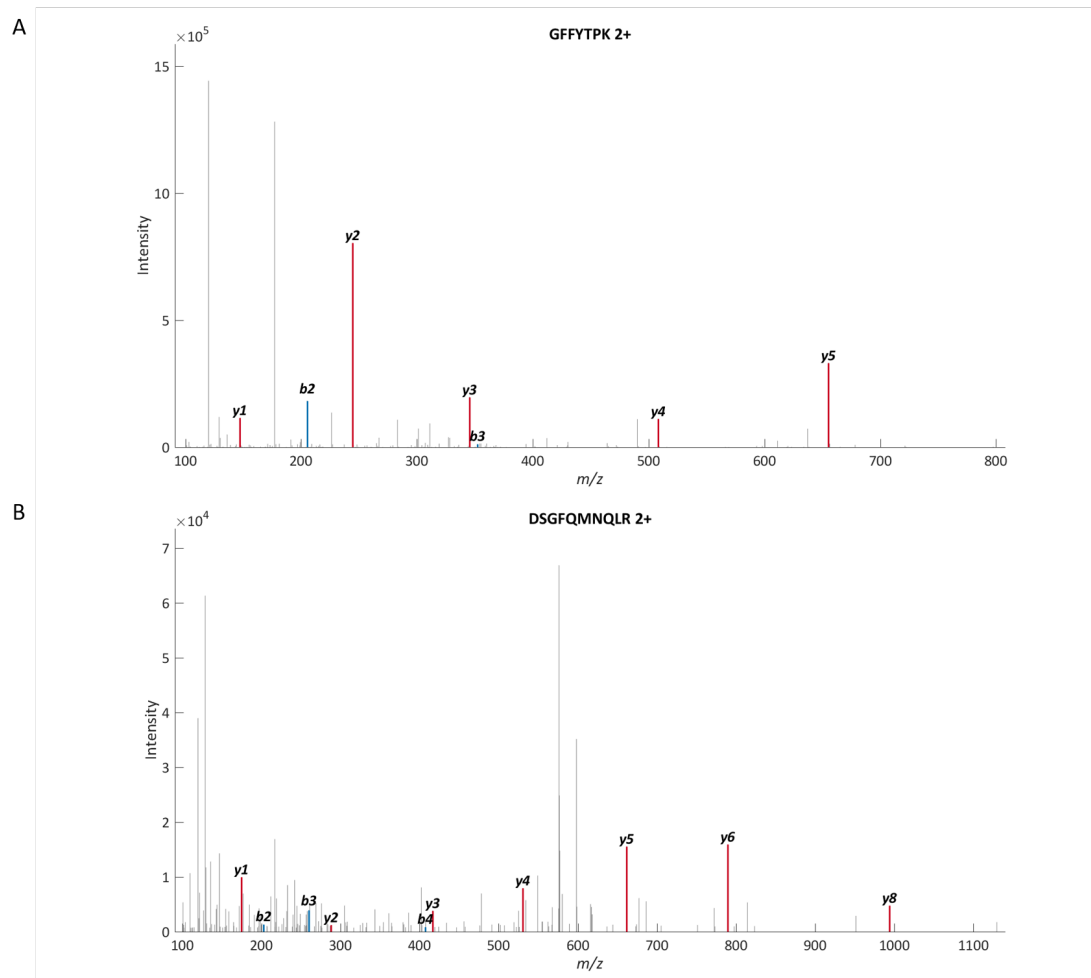
Insulin quantification was found to result in similar values at both MS<sup>1</sup> and MS<sup>2</sup> level in terms of accuracy and variation between replicates (Table 3.9). For transferrin, the difference between MS<sup>1</sup> and MS<sup>2</sup> was striking. CVs were substantially improved when MS<sup>2</sup> was used for quantification (most intense fragment ion). In general, MS<sup>1</sup> and MS<sup>2</sup> were found to perform similar for LC-MS/MS [238]. The potential downside of MS<sup>1</sup>-based quantification is stochastic sampling of precursor ions and therefore result in missing MS/MS spectra for quantification [238]. Therefore, a DIA strategy with subsequent spectral library searching would be more suitable for reproducible identification and quantification of proteins [214].

However, more attention has to be devoted to the accuracy of the quantification. For both insulin and transferrin, the accuracy (percentage difference between actual and calculated difference in concentration) was poor. Further investigation is necessary to find out whether the accuracy can be improved. This can be done using normalisation strategies which account for technical variation [133–136, 239–241]. Since there is no consensus in the field of MS-based proteomics on a golden standard quantification strategy [242], various MS-based quantification strategies are still being used,

potentially introducing subjectivity towards explaining biological phenomena.



**Figure 3.17:** Annotated MS<sup>1</sup> spectrum for tryptic peptides generated from transferrin. The arrow indicates the most intense peptide (DSGFQMNQLR), which is shown in detail in the figure insert



**Figure 3.18:** Fragmentation spectra used for MS<sup>2</sup> quantification. (A) Most intense peptide for insulin: GFFYTPK. (B) Most intense peptide for transferrin: DSGFQMNQLR

**Table 3.9:** Comparison of quantification strategies for insulin and transferrin

MS <sup>1</sup>				MS <sup>2</sup>		
Protein	CV 1x	CV 0.5x	Fold change	CV 1x	CV 0.5x	Fold change
Insulin	0.3%	0.4%	2.76 (38%)	1.06%	1.1%	2.86 (43%)
Transferrin	21.8%	8.6%	0.91 (55%)	1.5%	3.5%	1.02 (49%)

### 3.5 CONCLUSION

In this chapter, parameters for LESA-MS/MS analysis and conditions for were optimised. The use of ACN/H<sub>2</sub>O/DMSO/FA 50:50:1:0.1 (v/v/v/v) resulted in the most repeatable extraction of peptides from the surface. The optimal solvent volume for a stable spray was found to be at least 2.5 µL. Furthermore, repeated dispensing/aspiration cycles increased the variability in surface extraction due to loss of solvent and/or an increased air gap at the end of the tip. LESA solvent and ionisation parameters were initially optimised on Permanox<sup>TM</sup> cell culture slides and later adjusted for the Droplet Microarray.

Best performing search algorithms were X!Tandem, OMSSA, and MS-GF+ in terms of unique peptide identification and processing time. In terms of digestion, it was found that *in situ* digestion at 37°C for 4 hours using 0.05 sequencing grade trypsin in DMSO/100 mM AmBic (1:9 v/v) provided the most reproducible results. Further, it was found that a targeted acquisition mode (DIA) results in more reproducible identification of proteins than in DDA mode. In addition, MS<sup>1</sup> and MS<sup>2</sup> quantification perform similar in terms of repeatability and reproducibility. Besides applying this strategy for quantitative analysis of proteins adsorbed on biomaterial surfaces, this strategy can be used as a starting point for other quantitative screening studies of various proteins using LESA-MS/MS.

## CHAPTER 4

# **High throughput screening of acrylate and methacrylate homopolymers for examining adsorption of culture medium proteins in relation to cell response of human induced pluripotent stem cells (hiPSCs)**

### **4.1 INTRODUCTION**

**C**ELL-INSTRUCTIVE materials have a great potential in the manufacture of cells, as medical devices and in cell therapies. Human induced pluripotent stem cells (hiPSCs) are of great interest in regenerative medicine due to their ability to differentiate into any cell type in the three germ layers. However, more than a billion cells are required for each intervention. To date, the move to industrial-scale cell culture has

been limited by the incompatibility of biological growth substrates due to costs and large batch-to-batch variation. Therefore, research has focused on the discovery of synthetic growth substrates to facilitate reproducible and low-cost synthetic substrates for stem cell culture factories [243, 244].

The use of material microarrays allows biomaterials discovery to be progressed in a high throughput manner [14]. Material microarrays consists of numerous unique materials (>500 [24]) printed onto a modified substrate in order to perform large scale screening for a desired cell-instructive property, e.g. pluripotency maintenance of human induced pluripotent stem cells (hiPSCs) [33]. The use of high throughput materials screening strategies has led to the discovery of bacterial-resistant coatings and synthetic cell culture substrates [33, 245].

It is well established that pre-coating of synthetic polymers with extracellular matrix components increases cell attachment [246]. Therefore, it suggests that cells do not attach directly to the synthetic polymers, but synthetic polymers require a biological coating to allow cells to attach. So far, the missing part in biomaterials discovery has been the ability to qualitatively and quantitatively assess this biological coating adsorbed to the materials in high throughput to determine its role in the cell-instructive mechanism [66]. Hammad *et al.* [67] extracted surface-bound proteins from synthetic surfaces incubated with mouse embryonic fibroblast conditioned medium (MEF-CM) and used mass spectrometry-based proteomics [76] for identification. They have shown that pre-adsorption of certain cell culture medium proteins on plasma-etched tissue culture polystyrene (PE-TCPS) induces increased stem cell attachment.

To make quantitative screening compatible with the material microarrays, a strategy has been developed in the previous chapter. This strategy will now be employed to perform high throughput quantitative proteomics to investigate the relation between protein adsorption and hiPSC attachment and pluripotency maintenance on synthetic polymer surfaces. This information could provide fundamental understanding of the mechanism of cell-instructive polymers and will further assist in rationalisation and development of improved polymer structures for reproducible and cost-effective stem



cell culture factories.

## 4.2 MATERIALS & METHODS

### 4.3 REAGENTS & CHEMICALS

IPA (HPLC grade), DMF (99.8%), DMSO ( $\geq 99.9\%$ ), acetic acid (Optima<sup>TM</sup> LC-MS grade), and FA (Optima<sup>TM</sup> LC-MS grade) were acquired from Fisher Scientific (Loughborough, UK). DMSO (LC-MS grade; Pierce<sup>TM</sup>), Essential 8<sup>TM</sup> 50x supplement, 4% (v/v) FBS, Alexa500 rabbit antibody, Nanog antibody, and DMEM/F12 were purchased from Thermo Scientific. ACN (CHROMASOLV<sup>®</sup> LC-MS grade) and H<sub>2</sub>O (CHROMASOLV<sup>®</sup> LC-MS grade) were purchased from Honeywell (Seelze, Germany). Monomers (Appendix A) were acquired from Sigma-Aldrich and Polymer Sciences in the highest purity available. Sequencing grade trypsin was obtained from Promega (Southampton, UK). 2-DMPA (99%), recombinant human insulin ( $\geq 98\%$ ), and human *holo*-transferrin ( $\geq 97\%$ ) were acquired from Sigma-Aldrich (Gillingham, UK). Droplet Microarrays were provided by Aquarray (Karlsruhe, Germany). FGF-2 ( $>95\%$ ) and TGF- $\beta$ 1 ( $>97\%$ ) were obtained from R&D Systems (Abingdon, UK). ROCK inhibitor Y-27632 hydrochloride salt (ROCKi) was purchased from STEMCELL Technologies (Vancouver, Canada).

#### 4.3.1 ARRAY PRINTING

All printing optimisation was done on a Biodot XYZ3000 contact printer (Biodot, Irvine, CA) equipped with one 946PM6B steel pin (Arrayit, Sunnyval, CA). A selection of monomers and organic solvents was used to assess the printing performance. Each monomer solution was subjected to *in situ* polymerisation using a long-wave UV light. Printing was performed in an argon-rich environment (O<sub>2</sub> <2000 ppm). During optimisation the relative humidity was altered to determine the optimal printing condition. Printed arrays were dried for seven days in a vacuum oven (Thermo Scientific; 35°C , <50 mTorr) to extract residual solvent and after stored at room temperature.

### 4.3.2 TOF-SIMS

TOF-SIMS analysis was performed on a TOFSIMS IV (IONTOF GmbH, Münster, DE) equipped with a  $\text{Bi}_3^+$  primary ion source. The field of view was scanned in stage scan raster mode with a pixel density of 12.5 pixels per millimeter. Both positive and negative secondary ion spectra were collected within one scan per polarity. Analysis spectra and ion images was done in SurfaceLab 7.0 (IONTOF GmbH, Münster, Germany).

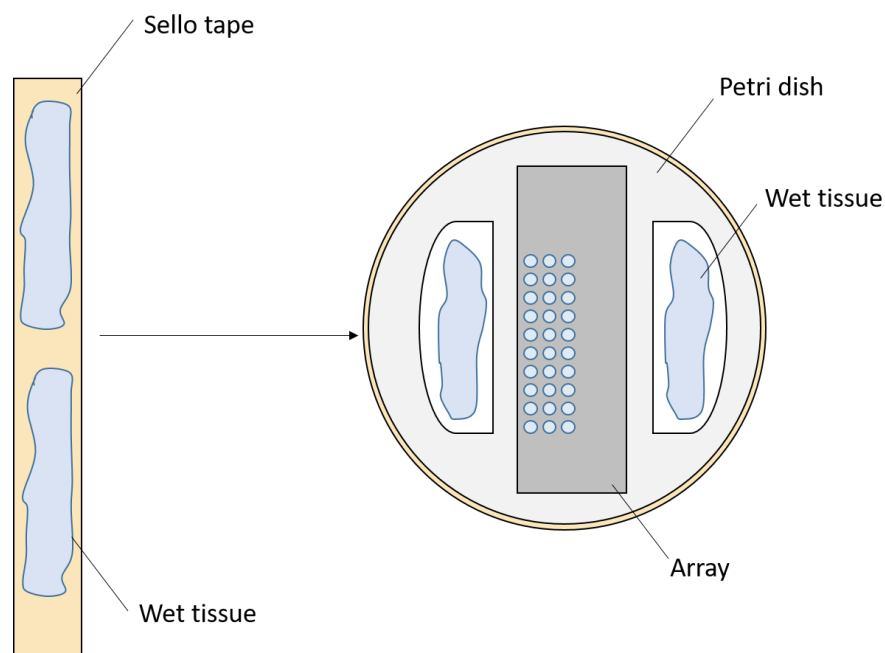
### 4.3.3 INCUBATING MICROARRAYS IN ESSENTIAL 8™

Printed arrays were incubated for 24 hours in deionised  $\text{H}_2\text{O}$  (18.2 M $\Omega$ ; Elga PureLab, High Wycombe, UK) at room temperature and after allowed to dry under ambient conditions. Next, a stock solution (10 mM) for ROCK inhibitor Y-27632 was prepared in deionised water. Thereafter, a clean 50 mL Falcon tube was filled with 49 mL DMEM/F12 and 1 mL Essential 8™ 50x supplement to which ROCK inhibitor Y-27632 was to reach a final concentration of 10  $\mu\text{M}$ . The array was then placed in the culture medium and incubated for 1 hour at 37 °C . After the incubation period, the array was dipped for 10 seconds in deionised  $\text{H}_2\text{O}$  three times and subsequently allowed to dry under ambient conditions. Directly after, *in situ* digestion was carried out.

### 4.3.4 *In situ* DIGESTION

A stock solution of sequencing grade trypsin (0.5 ) was prepared in 50 mM acetic acid and stored in 10  $\mu\text{L}$ -aliquots at -80°C until further use. Prior to application, the aliquot was thawed at room temperature and diluted with 90  $\mu\text{L}$  of DMSO/100 mM AmBic (1:9 v/v). The digestion solution was dispensed on the array using the rolling droplet technique. The array was subsequently placed in humidified chamber and incubated for four hours at 37°C . The humidity chamber was created in a Petri dish by placing wet tissue alongside the array and sealing the lid with wet tissue fixed on Sello tape (Figure 4.1). After incubation, the array was placed in a vacuum oven (Thermo Scientific; <50 mTorr) to extract the digestion solution. The array was subsequently stored

at -20°C until further analysis.



**Figure 4.1:** Schematic representation of the humidity chamber used for *in situ* digestion

#### 4.3.5 LESA-MS/MS FOR SPECTRAL LIBRARY GENERATION

Essential 8<sup>TM</sup> 50x was diluted to 25x and 10x in 100 mM AmBic. 50x, 25x, and 10x solutions were introduced on the Droplet Microarray using the rolling droplet technique and allowed to dry under ambient conditions. Thereafter, proteins were subjected to *in situ* digestion as described in Section 4.3.4 and LESA-MS/MS analysis.

LESA-MS/MS analysis was performed on a TriVersa Nanomate (Advion Biosciences, Ithaca, NY) coupled to a Q Exactive plus mass spectrometer. The extraction solvent consisted of ACN/H<sub>2</sub>O /DMSO/FA 50:50:1:0.1 (v/v/v/v). A total of 3 µL was aspirated from the solvent reservoir from which 1.5 µL was dispensed on a protein digest. After 15 seconds, 2.0 µL was aspirated back into the tip and infused into the mass spectrometer using a voltage of +1.6 kV with 0.4 psi N<sub>2</sub> back pressure. The mass spectrometer was operated in DDA mode. The isolation window was set to 1 *m/z*. The NCE was set to 27, the resolution was set to 17,500 at *m/z* 200, the AGC target was set

to  $1 \times 10^5$  with a maximum ion injection time of 250 ms, and the first mass (minimum  $m/z$ ) was fixed at  $m/z$  100. After every 20  $MS^2$  scans, a  $MS^1$  scan was acquired over the range  $m/z$  400-1200. The resolution was set to 140,000 at  $m/z$  200, the AGC target to  $3 \times 10^6$  and the maximum ion injection time to 200 ms.

#### 4.3.6 LESA-MS/MS FOR HIGH THROUGHPUT PROTEIN QUANTIFICATION ON POLYMER MICROARRAYS

An inclusion list was generated for all peptide ions present in the spectral library using XCalibur 2.1 (Thermo Scientific, San Jose, CA). LESA-MS/MS analysis was performed on a TriVersa Nanomate coupled to a Q Exactive plus mass spectrometer. The extraction solvent consisted of ACN/H<sub>2</sub>O /DMSO/FA 50:50:1:0.1 (v/v/v/v). A total of 3  $\mu$ L was aspirated from the solvent reservoir from which 1.5  $\mu$ L was dispensed on a protein digest. After 15 seconds, 2.0 L was aspirated back into the tip and infused into the mass spectrometer using a voltage of +1.6 kV with 0.4 psi N<sub>2</sub> back pressure. The mass spectrometer was operated in DIA mode. The isolation window was set to 1  $m/z$ . The NCE was set to 27, the resolution was set to 35,000 at  $m/z$  200, the AGC target was set to  $1 \times 10^6$  with a maximum ion injection time of 250 ms, and the first mass (minimum  $m/z$ ) was fixed at  $m/z$  100. After every 10  $MS^2$  scans, a  $MS^1$  scan was acquired over the range  $m/z$  400-900. The resolution was set to 140,000 at  $m/z$  200, the AGC target to  $3 \times 10^6$  and the maximum ion injection time to 200 ms. The total analysis time was 1 minute per polymer spot.

#### 4.3.7 PROTEIN IDENTIFICATION & QUANTIFICATION

The LESAProteomics pipeline was used to perform sequence database searching for identification of proteins and spectral library generation. MS/MS files were subjected against a customised FASTA file containing the amino acid sequences of the recombinant human proteins and the yeast proteome to serve as decoys (Table 4.1). Decoy sequences are used to estimate the FDR and validate peptide identification. The FDR was set to 5% (minimum score 95). X!Tandem [104], MS-GF+ [108], and OMSSA [103]

were used as search algorithms. Identified peptides per protein were used to extract precursor intensities from averaged MS<sup>1</sup> spectra. The peptide intensities per protein were summed and log-transformed.

**Table 4.1:** Amino acid sequences for target proteins in FASTA format

Protein	Sequence in FASTA format			
Insulin	GIVEQCCTS	ICSLYQLEN	YCNFVNQHL	CGSHLVEAL
	YLVCGERGF	FYTPKT		
Transferrin	VPDKTVRWC	AVSEHEATK	CQSFRDHMK	SVIPSDGPS
	VACVKKASY	LDCIRAI	AA NEADAVTLD	AGLVYDAYL
	APNNLK-	PVV AEFYGS	KED PQTFYYAVA	VVKKDSGFQ
	MNQLRGKKS	CHT-	GLGRSA	GWNIPIGLL
	YCDLPEPRK	PLEKAVANF	FSGSCAPCA	DGTD
	FPQLC	QLCPGCGCS	TLNQYFGYS	GAFKCLKDG
	AG-	DVAFVKH	STIFENLAN	KARDQYEL
	LCLDNTRKP	VDEYKD-	CHL	AQVPSHTVV
	ARSMGGKED	LIWELLNQA	QEHFGKDKS	KE-
	FQLFSSP	HGKDLLFKD	SAHGFLKVP	PRMDAKMYL
	GYEYVTAIR	NLREGTCPE	APTDECKPV	KWCALSHHE
	RLKCDEWSV	NS-	VGKIECV	SAETTEDCI
	AKIMNGEAD	AMSLDGGFV	YIAGK-	CGLV
	PVLAENYNK	SDNCEDTPE	AGYFAIAVV	KKSASDLTW
	DNLKGKKSC	HTAVGRTAG	WNIPMGLLY	NKINHCRFD
	EFF-	SEGCAP	GSKKDSSLC	KLCMGSGLN
	LCEPNNKEG	YYGYT-	GAFR	CLVEKGDVA
	FVKHQTV	PQ NTGGKNPDP	WAKNL-	NEKD
	YELLCLDGT	RKPVEEYAN	CHLARAPNH	AVVTRKDKE
	ACVHKILRQ	QQHLFGSNV	TDCSGNFCL	FRSETKD
	LL	FRD-	DTVCLA	KLHDRNTYE
	KYLGE	EYVK	AVGNLRKCS	TSSLLEACT
	FRRP			

**Table 4.1** (continued)

Protein	Sequence in FASTA format			
FGF-2	AAGSITTLPA	LPEDGGSGAF	PPGHFKDPKR	LYCKNGGFFL
	RIHPDGRVDG	VREKSDPHIK	LQLQAEERGV	VSIGVVCANR
	YLAMKEDGRL	LASKCVTDEC	FFFERLESNN	YNTYRSRKYT
	SWYVALKRTG	QYKLGSKTGP	GQKAILFLPM	SAKS
TGF- $\beta$ 1	ALDTNYCFSS	TEKNCCVRQL	YIDFRKDLGW	KWIHEPKGYH
	ANFCLGPCPY	IWSLDTQYSK	VLALYNQHNP	GASAAPCCVP
	QALEPLPIVY	YVGRKPKVEQ	LSNMIVRSCK	CS

Amino acid sequences for recombinant proteins retrieved from [247]

#### 4.3.8 MICROARRAY SCREENING OF HUMAN INDUCED PLURIPOTENT STEM CELLS

Microarray data were obtained from Nasir *et al.* [59]. Briefly,  $0.75 \times 10^6$  hIPSCs (cell line: Rebl-PAT) were seeded in Essential 8<sup>TM</sup> medium supplemented with 10 $\mu$ M ROCKi onto a polymer microarray and cultured for 24 hours at 37°C and 5% CO<sub>2</sub>. Cells were fixated using 4% (v/v) PFA for 20 minutes at room temperature in a dark environment. Arrays were then immunostained for OCT4 expression (pluripotency marker) and counterstained with DAPI (cell nucleus marker). Images were acquired using an automated fluorescence microscope (IMSTAR). Counting of total cell number and number of pluripotent stem cells was achieved through CellProfiler v2.2.0 [61].

#### 4.3.9 DATA CLEANING & STATISTICAL ANALYSIS

Unsuccessful printed polymers (not observed in ToF-SIMS ion images) were removed from the data. Summed peptide intensities were subsequently quantile normalised. In quantile normalisation, data from each replicate are forced into the same distribution by replacing each data point with the mean of the corresponding quantile in order

to reduce technical variation [231] The datasets (protein and cell data) were subjected to Thompson's tau test for identification of outliers [248]. The outlier region ( $RR$ ) is calculated according to Equation 4.1, in which  $N$  is the number of repeats and  $t_{\alpha/2}$  is the critical value derived from Student's t-distribution at significance level  $\alpha$ .

$$RR = \frac{t_{\alpha/2}(N-1)}{\sqrt{N}\sqrt{N-2+t_{\alpha/2}^2}} \quad (4.1)$$

Identified outliers in the protein data were replaced using the knn algorithm. Polymers with zero values for more than two replicates were removed. The algorithm calculates the a weighted average of the  $k$  nearest neighbouring values. The value of  $k$  was set to 11 [249] and Euclidean distance was used as distance function (Equation 4.2) which describes the distance ( $d$ ) as a direct line between two data points ( $p$  and  $q$ ) using the Pythagoras theorem.

$$d = \sqrt{\sum ((p_i - q_i)^2)} \quad (4.2)$$

Outliers in cell data were removed and not replaced. Protein data was checked for normality (data distribution) using Lilliefort's test. Statistical comparison of differences in protein adsorption was then performed using either parametric (ANOVA) or non-parametric statistics (Kruskal-Wallis) for multiple group comparison. The remaining data was subjected to regression analysis to investigate the relation between protein adsorption (independent variable) and cell count (dependent variable). Protein data was transformed onto a relative scale by dividing all by the maximum summed peptide intensity for each replicate [66]. Then a sigmoid curve was fitted to data, for which the general form is shown in Equation 4.3. A sigmoid curve indicates that a process reaches saturation in the dependent variable when the independent variable is increasing. This was observed by Mei *et al.* [66], i.e. protein adsorption did not further increase from a certain number of attached cells.

$$y(x) = \frac{1}{(1 + e^{-x})} \quad (4.3)$$

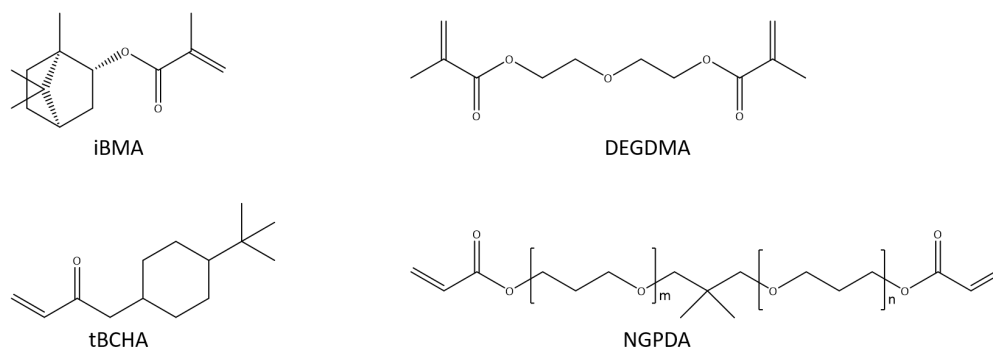
Further analysis was done to identify chemistry-related protein adsorption using PLS regression with the protein quantification data (response) and ToF-SIMS ions (variables). First, a PLS model was created to capture at least 90% of the variance in the data. VIP scores were calculated and a VIP score  $> 1.5$  was considered to be important to the PLS model and included to build a new predictive model. A new PLS model was then built. The optimal number of PLS components was selected based on the minimal mean square error of the prediction, which was calculated using 5-fold cross validation with 50 Monte-Carlo repetitions. After, PLS models were validated using a permutation test [250]. From the PLS model, the ions  $< m/z$  100 were assessed for relations between polymer chemistry and protein adsorption.

## 4.4 RESULTS & DISCUSSION

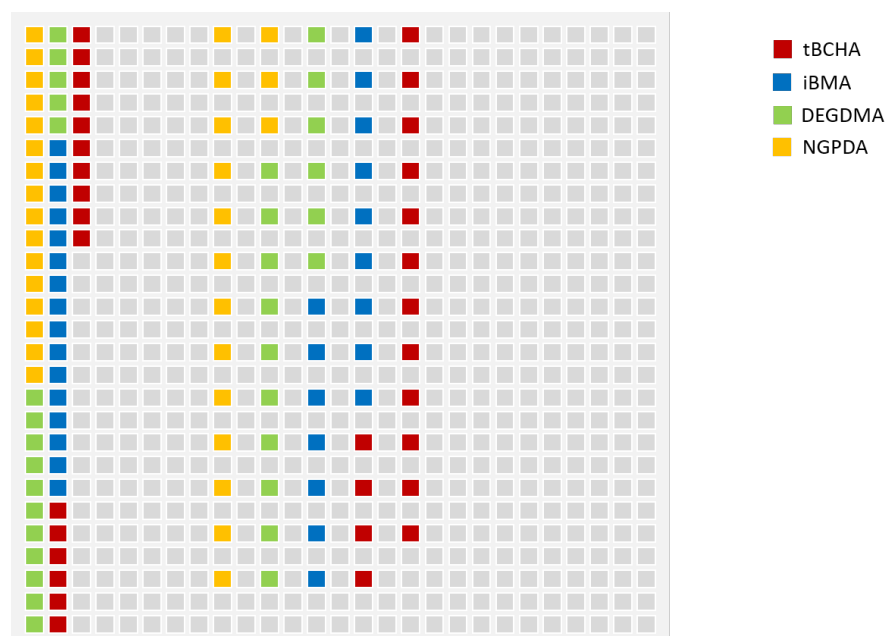
### 4.4.1 OPTIMISATION OF PRINTING PARAMETERS

In the first round of optimisation for printing monomer solutions on the Droplet Microarray ( $250\text{ }\mu\text{m} \times 250\text{ }\mu\text{m}$  square superhydrophilic features), four monomers (isobornyl methacrylate (iBMA), diethylene glycol dimethacrylate (DEGDMA), *tert*-butyl cyclohexylacrylate (tBCHA), and neopentyl glycol propoxylate diacrylate (NGPDA); Figure 4.2) were chosen to determine whether the Droplet Microarray is suitable to be used as printing substrate. These four monomers are chemically and structurally different and would therefore provide sufficient initial information about the suitability of the DMA as printing substrate. All monomers were prepared as 50% (v/v) solution in DMF after which a 3% (w/v) 2-DMPA solution was added in a 2:1 (v/v) ratio. Printing of the monomers was done using a single contact print onto a superhydrophilic spot. Each monomer solutions was printed for two rounds with 16 repeats. In the first round, the solution was dispensed on 16 consecutive spots. In the second round, monomer were printed on every other spot (Figure 4.3). This was to determine whether the printing solution merges to adjacent spots and would therefore cause cross contamination.





**Figure 4.2:** Structures for selected monomers for testing the Droplet Microarray as printing substrate



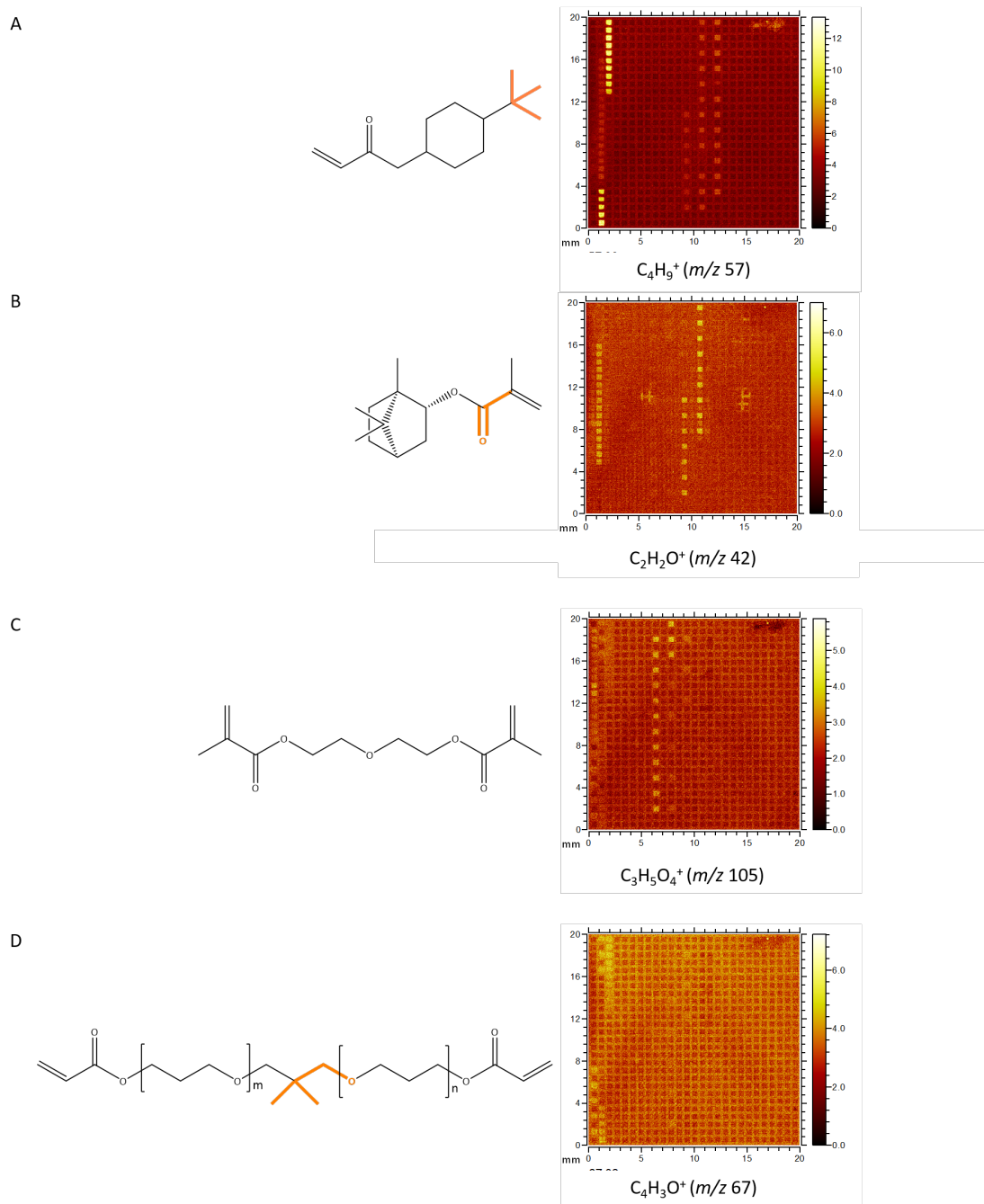
**Figure 4.3:** Schematic representation of print layout for testing the Droplet Microarray as printing substrate

TOF-SIMS analysis was performed on the printed arrays to investigate whether the monomers were successfully printed and polymerised on the superhydrophilic features of the Droplet Microarray. SurfaceLab was used to extract ion images representing the four different chemical structures. An in-house library of TOF-SIMS ions was used to find representative ions for each polymer. The ion images show successful and repeatable printing of tBCHA and iBMA on the Droplet Microarray (Figure 4.4A-B). For NGPDA and DEGDMA were also detected through TOF-SIMS analysis, though, not for all repeats (Figure 4.4C-D). This can be either related to the solution not being

## CHAPTER 4

dispensed onto the surface or that the *in situ* polymerisation was not successful. A reason for decreased printing efficiency could be related to the higher viscosity of NG-PDA and DEGDMA. Highly viscous solutions often cause problems in high throughput printing [251]. From this could be derived that the printing efficiency of viscous monomers could be improved by choosing a different solvent system or by reducing the monomer concentration.

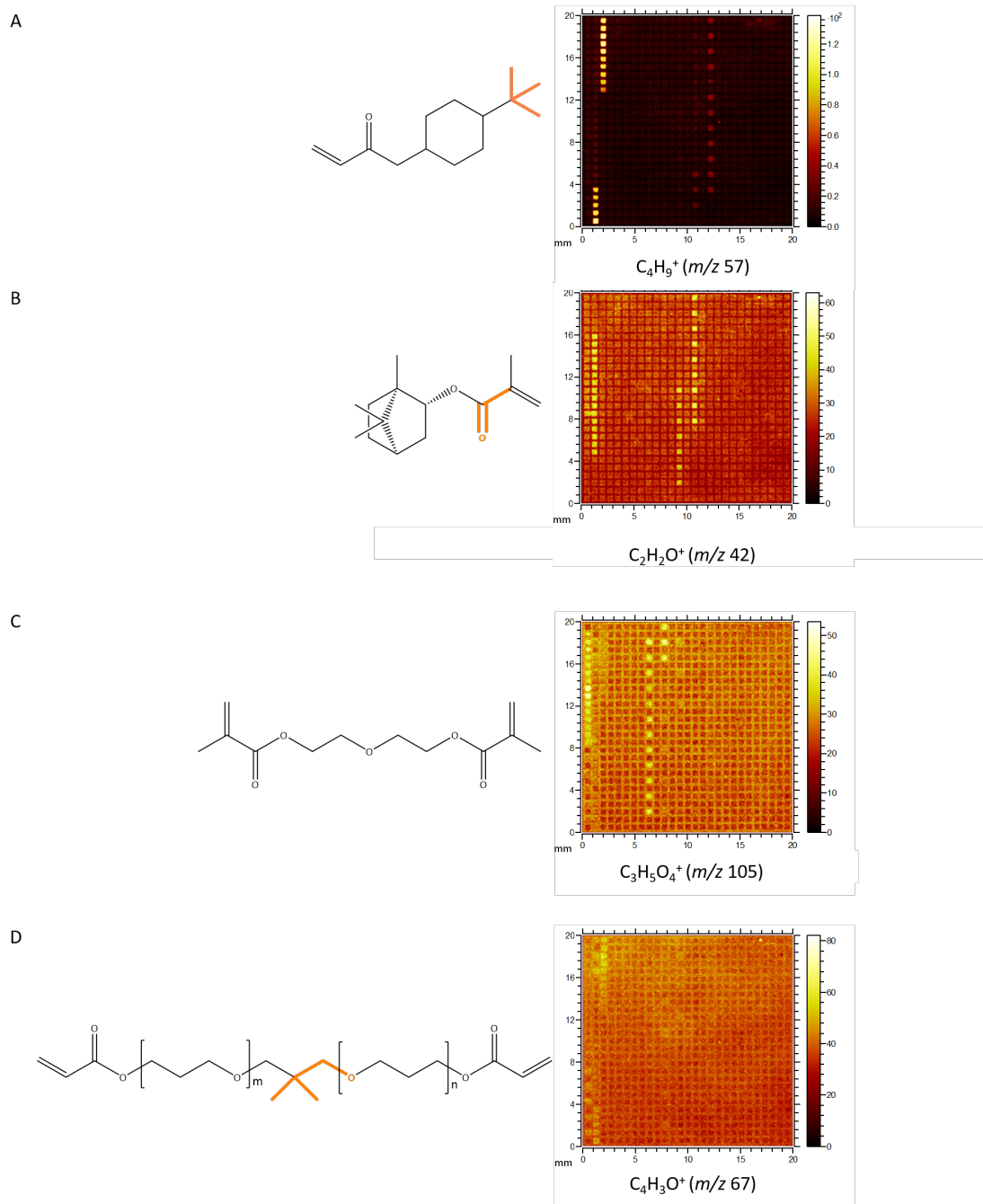
## CHAPTER 4



**Figure 4.4:** Representative TOF-SIMS ion images for the selected monomers before seven days of incubation in deionised H<sub>2</sub>O . (A) tBCHA (B) iBMA (C) DEGDMA (D) NGPDA. Chemical moieties detected by TOF-SIMS are highlighted in orange on the chemical structures

After TOF-SIMS analysis, the array was incubated in deionised H<sub>2</sub>O for seven days at room temperature. This was done to investigate the stability of the polymers on the Droplet Microarray when exposed to an aqueous environment. After incubation, the array was again analysed by TOF-SIMS. Using the same ions, it was confirmed that the successfully printed polymers were still present on the array (Figure 4.5A-D). When comparing the TIC normalised ion intensities before and after incubation, Student's *t*-test revealed a significant difference in ion intensity for the selected polymers (Table 4.2). To investigate whether the reduced ion intensity can be attributed to variation in instrument response, the fluorine peak (F<sup>-</sup>; *m/z* 19) was used as a reference. The superhydrophobic border contains of a fluorinised poly(HEMA-*co*-EDMA) polymer. The TIC normalised intensity for F<sup>-</sup> was 0.116 and 0.121 before and after wash, respectively. This indicates that the difference in ion intensity was not a result of fluctuation in instrument performance. Since the array was not washed before TOF-SIMS analysis, the ion intensities could be partially attributed to residual unreacted residual monomers on the surface which were washed away during the incubation in deionised H<sub>2</sub>O. However, long-term exposure (>1 day) of array to solutions should potentially be avoided.

## CHAPTER 4



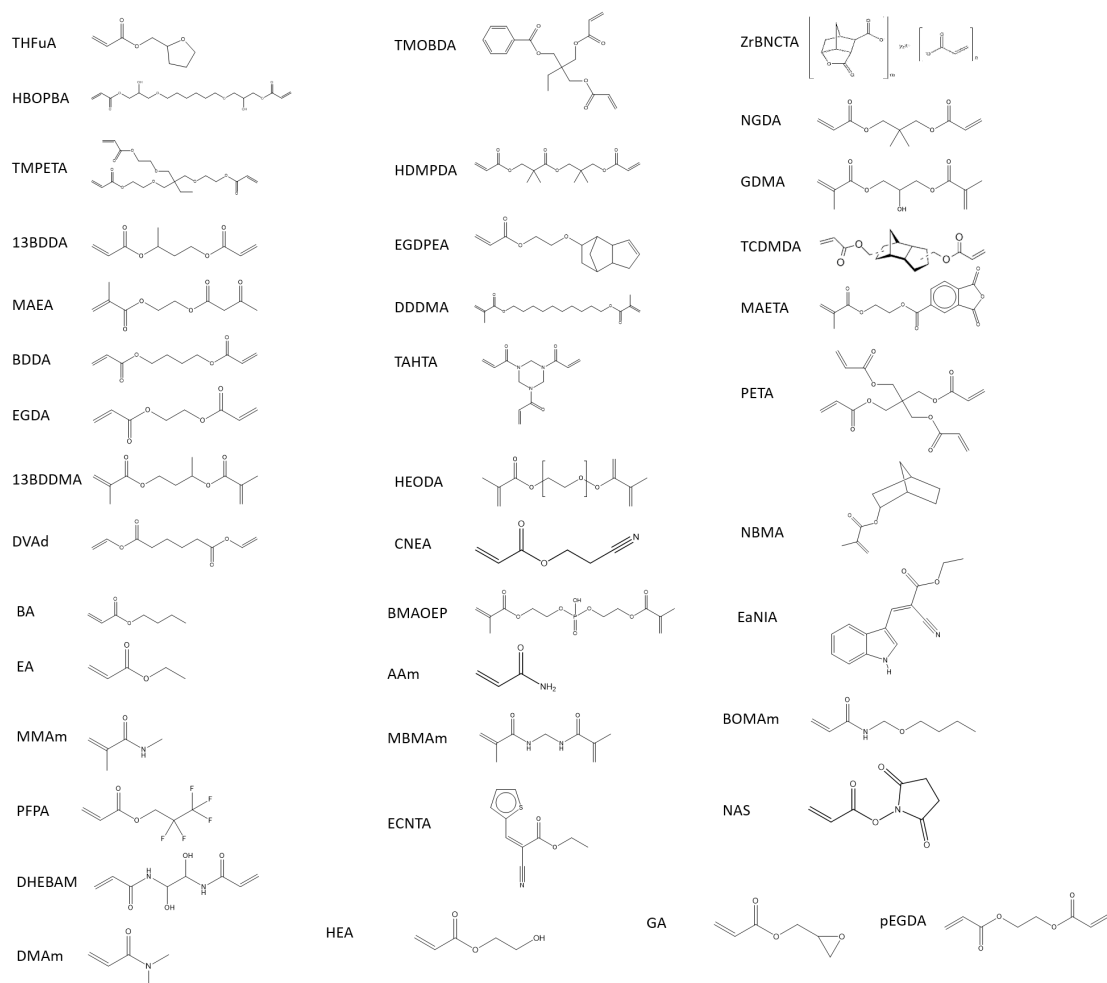
**Figure 4.5:** Representative TOF-SIMS ion images for the selected monomers after seven days of incubation in deionised  $\text{H}_2\text{O}$ . (A) tBCHA (B) iBMA (C) DEGDMA (D) NGPDA. Chemical moieties detected by TOF-SIMS are highlighted in orange on the chemical structures

**Table 4.2:** Comparison of TOF-SIMS ion intensities before and after wash

Polymer	Fold change	p-value
DEGDMA	1.32	$4.03 \times 10^{-5}$
NGPDA	1.27	0.01
iBMA	1.01	0.47
tBCHA	1.25	$1.29 \times 10^{-6}$

The previous Droplet Microarray format ( $500 \mu\text{m} \times 500 \mu\text{m}$ ) was not found to be compatible with the movement of the Nanomate robot, i.e. the spacing between spots is too small for the robot the move between adjacent superhydrophilic spots. Therefore, a format with a larger superhydrophilic spot size ( $\varnothing 1.414 \text{ mm}$ ) was used. The same printing program as in the first optimisation experiment was used, but now on a larger set of polymers to explore a larger chemical space (Figure 4.6). Also here, monomer solutions were prepared as 50% (v/v) or 50% (w/v) in DMF, where 2-DMPA was added to reach a concentration of 1% (w/v).

## CHAPTER 4



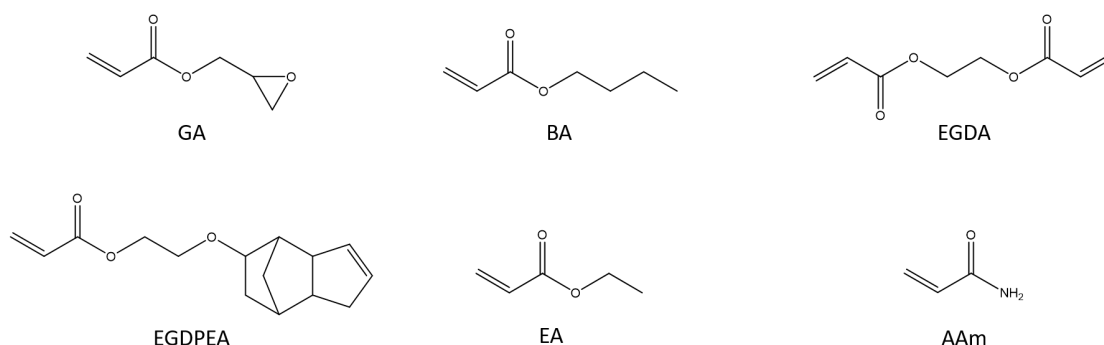
**Figure 4.6:** Structures used for second optimisation experiment

In total, 40 monomers were printed onto the superhydrophilic spots and *in situ* polymerised. Polymers were selected as representation of the available monomer library and further investigate the suitability of the Droplet Microarray for printing chemically-diverse monomers. Only 6 out of the 40 polymers were not detected using TOF-SIMS (Table 4.3). The structures of the monomer entities of the unsuccessfully printed polymers are shown in Figure 4.7. Five of those monomers were acrylate monomers and one acrylamide. The two other structures had a single ethylene glycol moiety. Though, poly(ethylene glycol) moieties were printed successfully indicating further optimisation (solvent system and printing parameters) is required for those monomers which have a single ethylene glycol.

**Table 4.3:** Overview of detected and not detected polymers using TOF-SIMS

Polymer	Detected	Polymer	Detected
THFuA	✓	TCDMDA	✓
HBOPBA	✓	MAETA	✓
TMPETA	✓	PETA	✓
13BDDA	✓	NBMA	✓
MAEA	✓	BA	✗
BDDA	✓	EA	✗
EGDA	✗	MMAm	✓
13BDDMA	✓	PFPA	✓
DVAd	✓	DHEBAM	✓
TMOBDA	✓	DMAm	✓
HDMPDA	✓	GA	✗
EGDPEA	✗	pEGDA	✓
DDDMA	✓	BMAOEP	✓
TAHTA	✓	AAm	✗
pEGDMA	✓	MBMAm	✓
HEODA	✓	ECNTA	✓
CNEA	✓	HEA	✓
ZrBNCTA	✓	BOMAm	✓
NGDA	✓	NAS	✓
GDMA	✓	EaNIA	✓

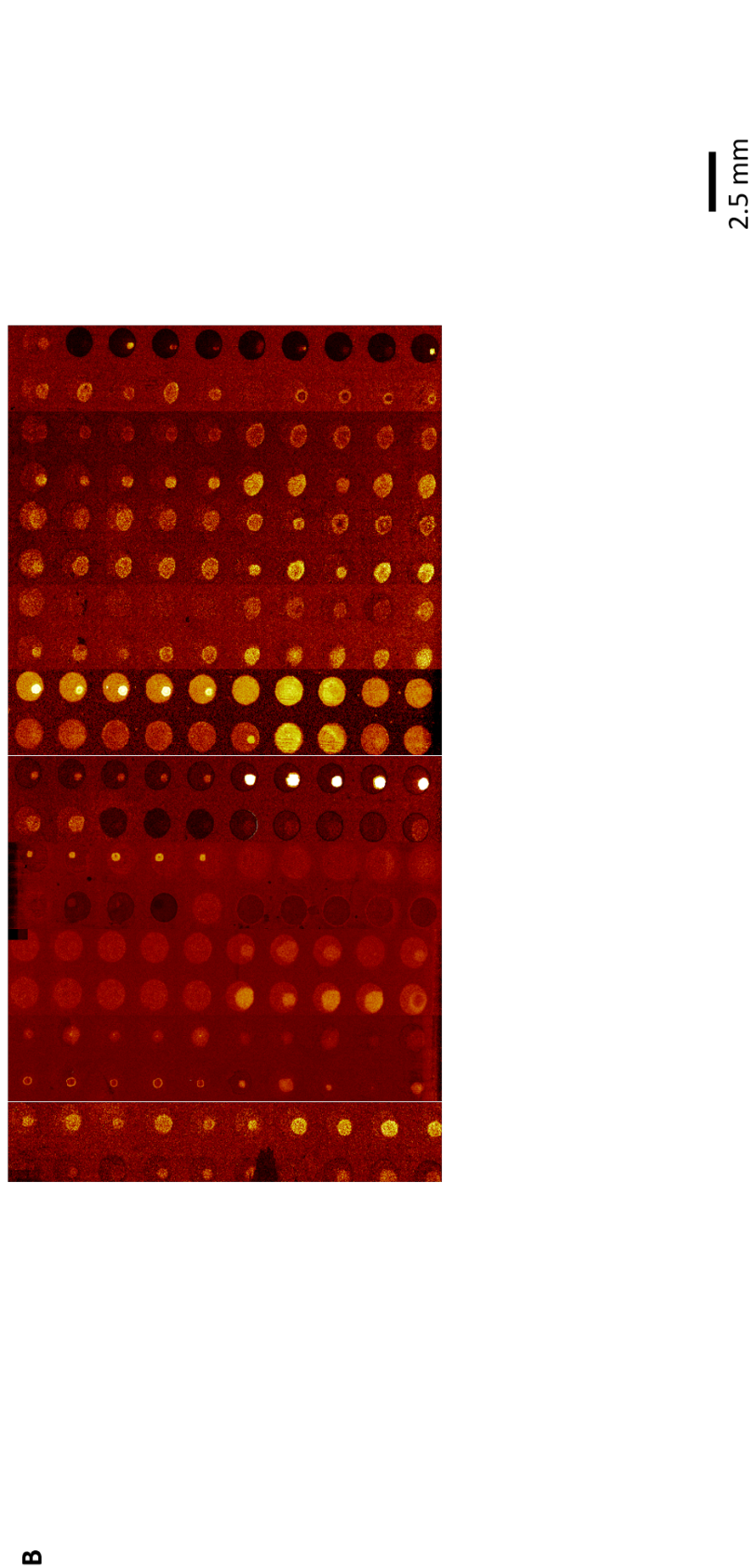




**Figure 4.7:** Chemical structures for unsuccessfully printed monomers. (A) glycidyl acrylate (GA), (B) butyl acrylate (BA), (C) ethylene glycol diacrylate (EGDA), (D) ethylene glycol dicyclopentyl ether acrylate (EGDPEA), (E) ethyl acrylate (EA), and (F) acrylamide (AAm)

The TOF-SIMS ion images were studied to further identify parameters requiring optimisation. In Figure 4.4.1, stitched TOF-SIMS ion images are shown for a representative ion for each polymer. From these ion images can be derived that the polymer spot is smaller than the substrate. This reveals that the standard printing parameters for material microarray production need to be further optimised to allow full spot coverage. To eliminate the possibility of the superhydrophilic polymer influencing the protein adsorption, the polymer of interest should cover the full area of the superhydrophilic spot. Increased volume of monomer solution onto the surface can be achieved through repetitive contact of the steel spin or by increasing the contact time. The contact time as well as the number of prints were altered to find the optimal condition of solvent dispensing. Three monomers (tricyclodecane-dimethanol diacrylate (TCDMDA), hexanediol ethoxylate diacrylate (HEODA), and methacryloyloxyethyl acetoacetate (MAEA)) were selected based on the observation that these monomer solution, among other, did not fully fill a spot. Monomer solutions (50% (v/v)) were prepared in DMF, DMSO, or ACN. Printing was performed using 75 contacts on the surface. ACN was found to be compatible with the Droplet Microarray (i.e. ACN is repelled by the superhydrophobic pattern) and permitting solubilisation of monomers, however, ACN evaporated too quickly (directly after contact with the surface) during the printing process preventing *in situ* polymerisation (Figure 4.8A). DMF was already found to be compatible in

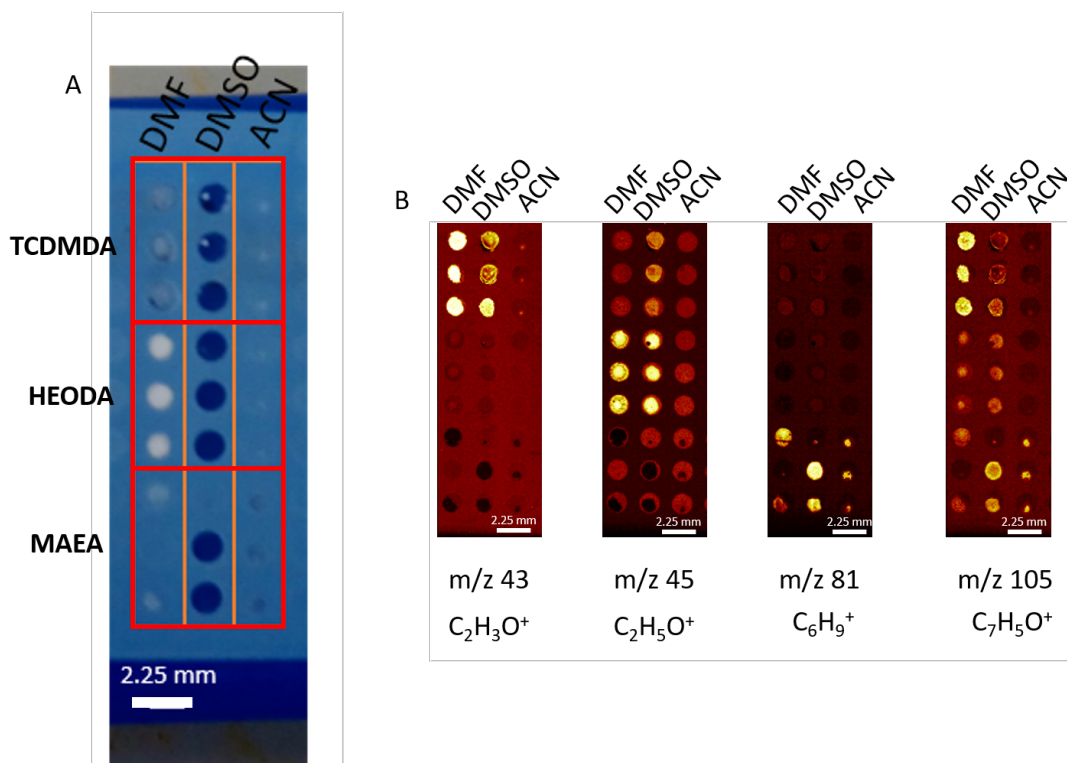
A	A	TCMDA	MAETA	PETA	NBMA	BA	EA	MMam	PFA	DHEBAM	DMam	GA	pEGDA	BMAOEP	AAm	MBMam	ECNTA	HEA	BOMam	NAS	EaNA
	B	TCMDA	MAETA	PETA	NBMA	BA	EA	MMam	PFA	DHEBAM	DMam	GA	pEGDA	BMAOEP	AAm	MBMam	ECNTA	HEA	BOMam	NAS	EaNA
	C	TCMDA	MAETA	PETA	NBMA	BA	EA	MMam	PFA	DHEBAM	DMam	GA	pEGDA	BMAOEP	AAm	MBMam	ECNTA	HEA	BOMam	NAS	EaNA
	D	TCMDA	MAETA	PETA	NBMA	BA	EA	MMam	PFA	DHEBAM	DMam	GA	pEGDA	BMAOEP	AAm	MBMam	ECNTA	HEA	BOMam	NAS	EaNA
	E	TCMDA	MAETA	PETA	NBMA	BA	EA	MMam	PFA	DHEBAM	DMam	GA	pEGDA	BMAOEP	AAm	MBMam	ECNTA	HEA	BOMam	NAS	EaNA
	F	THFUA	HBOPBA	TMPETA	138DDA	MAEA	BDDA	EGDA	138DDMA	DVAd	TMOBDA	HDMPOA	EGDPEA	DDDMA	TAHTA	pEGDMA	HEODA	CNEA	ZrBNCTA	NGDA	GDMA
	G	THFUA	HBOPBA	TMPETA	138DDA	MAEA	BDDA	EGDA	138DDMA	DVAd	TMOBDA	HDMPOA	EGDPEA	DDDMA	TAHTA	pEGDMA	HEODA	CNEA	ZrBNCTA	NGDA	GDMA
	H	THFUA	HBOPBA	TMPETA	138DDA	MAEA	BDDA	EGDA	138DDMA	DVAd	TMOBDA	HDMPOA	EGDPEA	DDDMA	TAHTA	pEGDMA	HEODA	CNEA	ZrBNCTA	NGDA	GDMA
	I	THFUA	HBOPBA	TMPETA	138DDA	MAEA	BDDA	EGDA	138DDMA	DVAd	TMOBDA	HDMPOA	EGDPEA	DDDMA	TAHTA	pEGDMA	HEODA	CNEA	ZrBNCTA	NGDA	GDMA
	J	THFUA	HBOPBA	TMPETA	138DDA	MAEA	BDDA	EGDA	138DDMA	DVAd	TMOBDA	HDMPOA	EGDPEA	DDDMA	TAHTA	pEGDMA	HEODA	CNEA	ZrBNCTA	NGDA	GDMA



**Figure 4.8:** (A) Printing layout for testing default printing parameters on a representative selection of monomers. (B) Stitched TOF-SIMS ion images for the selected polymers reveal poor polymer coverage on the superhydrophilic areas using default printing parameters for polymer microarrays. Images are a sum of the ions  $C_2H_3^+$ ,  $C_3H_3O^+$ ,  $C_2H_4O^+$ ,  $C_3H_5O_4^+$ ,  $CN^+$ . Representative ions were selected from an in-house TOF-SIMS library.

the first printing experiment. The solutions evenly distributed along the superhydrophilic area. However, due to longer exposure to DMF it was visually observed that the superhydrophobic border dissolved, i.e. the polymer layer disappeared. For DMSO, also good and even distribution along the spot was observed. DMSO is less volatile than DMF and ACN, as a result, the printed solutions are still present after the printing process was finished. For both ACN and DMF, the solvent already evaporated before the printing programme was finished (Figure 4.8A).

TOF-SIMS analysis (Figure 4.8B) revealed successful printing of MAEA and HEODA for solutions in 50% DMSO or DMF (v/v). Printing performance for 50% ACN (v/v) solutions was found to be poor due to the rapid solvent evaporation. Since, DMF tends to spread across the superhydrophobic pattern and could therefore cause spreading across multiple spots, DMSO was chosen as the most suitable solvent to prepare the monomer solutions.



**Figure 4.9:** Testing different solvent systems for printing monomer solutions onto the Droplet Microarray. (A) Photograph of printed array directly after completion of the printing program. (B) TOF-SIMS ion images for selected polymers for test printing. Representative ions were selected from an in-house TOF-SIMS library

Further testing was done on optimising the number of contacts and the length of contact time. It was visually observed that longer contact times were more efficient in terms of solvent dispensing than quick repetitive contacts. For more viscous monomer solutions, more repetitive contacts were found to benefit the dispensing. Therefore, the printer was set to dispense the monomer solutions five times with five seconds contact per spot.

The relative humidity was also investigated as parameter to improve the printing. Previous research has shown that increasing the relative humidity results in improved the release of the solvent from the pins [252]. At increased relative humidity (70%), a number of monomer solutions formed hydrogels resulting in printing failure. It is known that increased water content increases aggregation of molecules and could

therefore cause gel formation [253]. At reduced humidity (30%), the solutions tend to evaporate relatively fast and could therefore prevent *in situ* polymerisation. At the initially used humidity (50%) none of the problems found at higher or lower humidity were observed.

#### 4.4.2 BUILDING A SPECTRAL LIBRARY FOR PROTEIN IDENTIFICATION AND QUANTIFICATION

It is known that not all peptides from the same protein are representative for the relative amount of protein [237]. Therefore, a proper choice has to be made on which peptide to be used for protein quantification. Essential 8<sup>TM</sup> supplement was serially diluted at concentrations of 50 $\times$ , 25 $\times$ , and 10 $\times$ , respectively. The Essential 8<sup>TM</sup> dilutions were dispensed via the rolling droplet technique onto the Droplet Microarray (8 repeats), *in situ* digested and analysed with LESA-DIA-MS/MS. For each protein, the identified PSMs are shown in Table 4.4. Excellent sequence coverage was obtained selected PSMs for insulin (Figure 4.9), transferrin (Figure 4.10), FGF-2 (Figure 4.11), and TGF- $\beta$ 1 (Figure 4.12). PSMs per protein were saved as library in .MAT format using LESAProteomics. This library was used for identifying PSMs in the dataset acquired from the polymer arrays.

**Table 4.4:** Identified PSMs per Essential 8<sup>TM</sup> protein

Protein	Peptide	Charge
<b>Insulin</b>	GFFYTPK	1+
	GFFYTPK	2+
<b>Transferrin</b>	MYLGYEYVTAIR	2+
	HSTIFENLANK	2+
	YLGEYVK	2+
	KCSTSSLLEACTFR	2+
	SASDLTWDNLKGKK	3+
	SASDLTWDNLKGK	2+

Table 4.4 (continued)

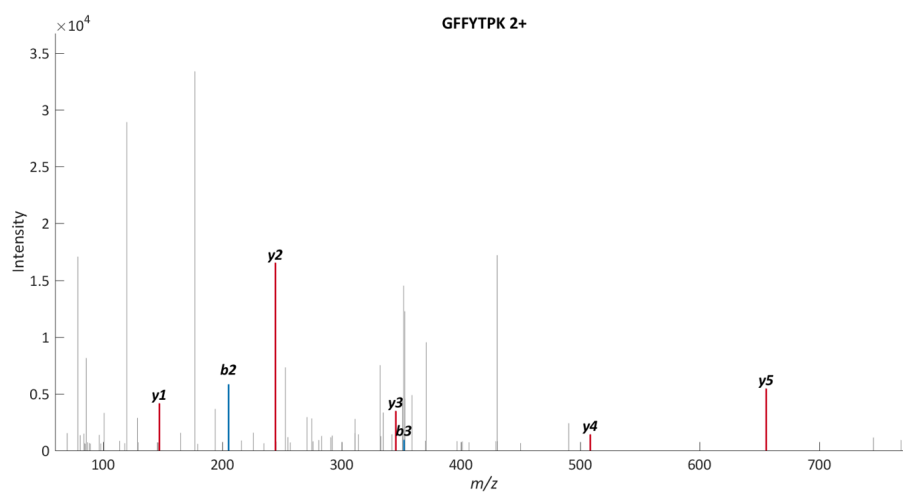
Protein	Peptide	Charge
	NTYEKYLGEELYVK	2+
	DSGFQMNQLR	2+
	SASDLTWDNLK	2+
	EGTCPEAPTDECKPVK	2+
	HQTVPQNTGGKNPDPWAK	3+
	KDSGFQMNQLR	2+
	FDEFFSEGCAPGSKK	2+
	DGAGDVAFVK	2+
	HQTVPQNTGGK	2+
	SKEFQLFSSPHGK	3+
	SVIPSDGPSVACVKK	2+
	EFQLFSSPHGK	2+
	DSGFQMNQLRGK	2+
	KSASDLTWDNLKGK	3+
	CSTSSLLEACTFRRP	2+
	ADRDQYELLCLDNTR	2+
	KSASDLTWDNLK	2+
	MYLGYEYVTAIRNLR	3+
	SMGGKEDLIWELLNQAQEHFGKDK	4+
	SMGGKEDLIWELLNQAQEHFGK	3+
	EGYYGYTGAFRCLVEKGDVAFVK	4+
	DQYELLCLDNTR	2+
	WCALSHHERLK	2+
	DCHLAQVPSHTVVAR	2+
	DQYELLCLDNTRK	2+
	DDTVCLAKLHDR	2+

Table 4.4 (continued)

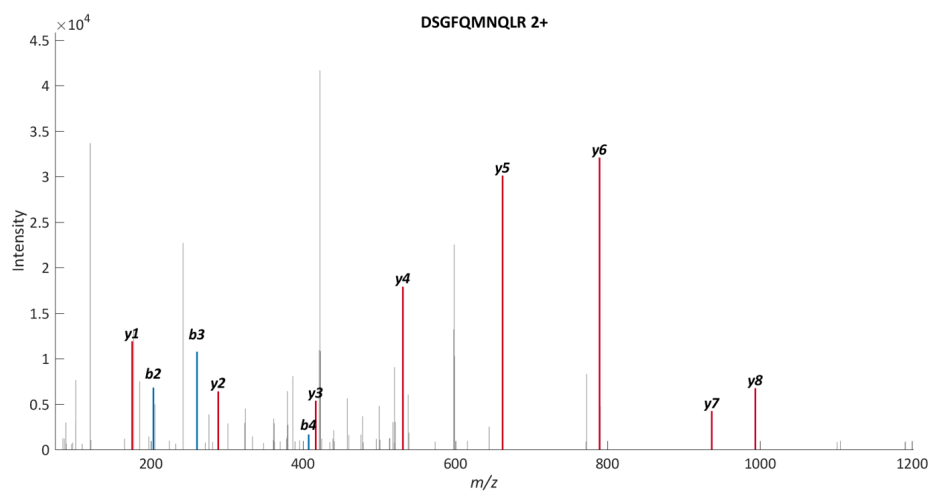
Protein	Peptide	Charge
	EGYYGYTGAFR	2+
	TAGWNIPMGLLYNK	2+
	APNHAVVTR	4+
	DSAHGFLKVPPRMDAK	3+
	WCALSHHER	1+
	YLGEYVKAVGNLR	2+
	DLLFRDDTVCLAK	2+
	DKEACVHKILR	2+
<b>TGF-<math>\beta</math>1</b>	ALDTNYCFSSTEKNCCVRQLYIDFR	3+
	NCCVRQLYIDFR	2+
	VLMVETHNEIYDKFK	2+
	KRIEAIRGQILSK	3+
	RIEAIR	1+
	RIEAIRGQILSK	2+
	YSNNSWRYLSNR	2+
	VEQLSNMIVRSCK	2+
	RALDTNYCFSSTEK	2+
	WIHEPK	2+
<b>FGF-2</b>	SRKYTSWYVALK	2+
	TRGRRTEERPSGSR	2+
	IHPDGRVDGVREK	2+
	SDPHIKLQLQAEER	2+
	GVCANRYLAMKEDGR	2+
	MVGVGGGDVEDVTPR	2+

**Table 4.4** (continued)

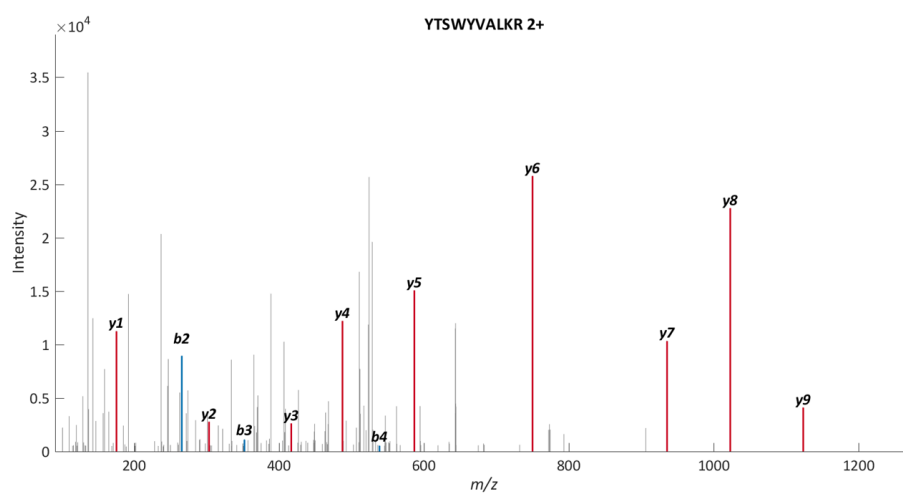
Protein	Peptide	Charge
	LESNNYNTYRSRK	2+
	EKSDPHIK	1+
	VGVGGGDVEDVTPRPGGCQISGR	3+

**Figure 4.10:** Fragmentation spectra for selected peptide GFFYTPK for identification of insulin

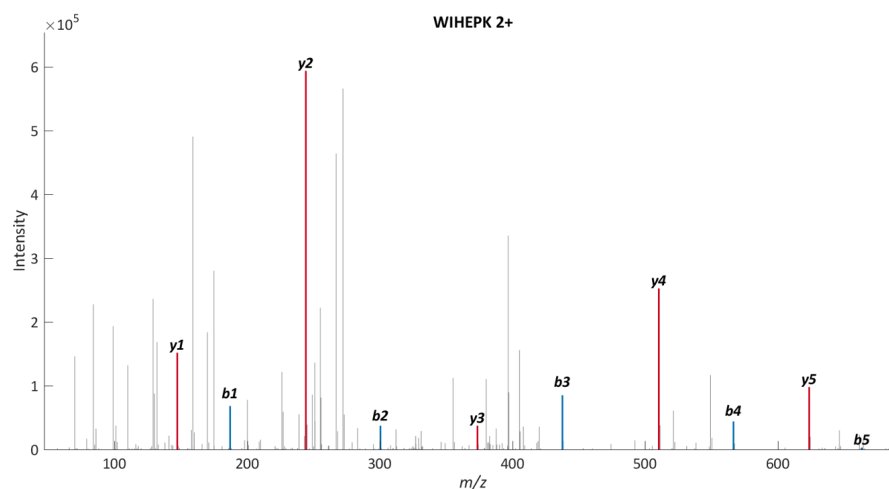




**Figure 4.11:** Fragmentation spectra for selected peptide DSGFQMNQLR for identification of transferrin.



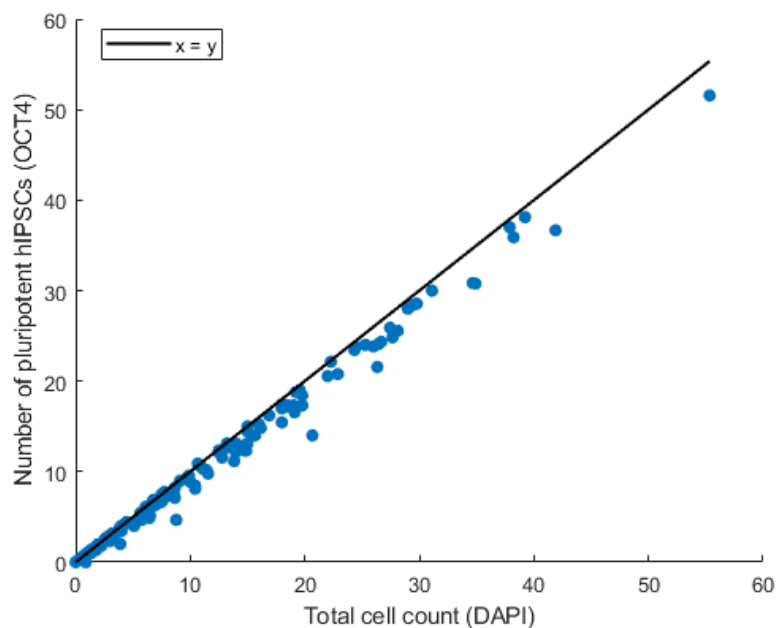
**Figure 4.12:** Fragmentation spectra for selected peptide YTSWYVALKR for identification of FGF-2



**Figure 4.13:** Fragmentation spectra for selected peptide WIHEPK for identification of TGF- $\beta$ 1 .

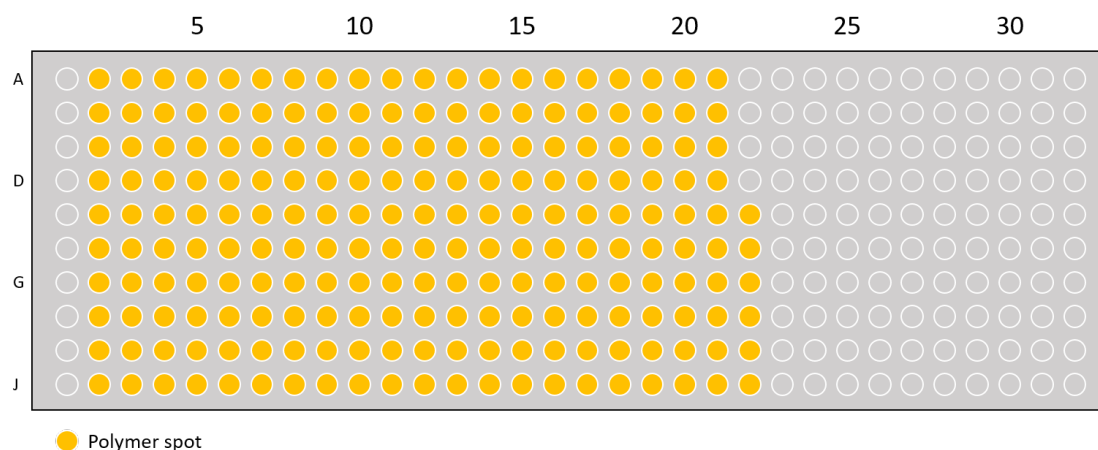
#### 4.4.3 ESSENTIAL 8<sup>TM</sup> PROTEIN DOES NOT CORRELATE WITH ATTACHMENT OF PLURIPOTENT STEM CELLS

With the peptides for quantification being selected, the quantification strategy was tested on a selection of polymers which were previously used for screening of human induced pluripotent stem cell (hiPSC) attachment and pluripotency maintenance [59]. ReBl PAT hiPSCs were cultured on polymer microarray by Dr Jordan Thorpe to identify suitable chemical structures as alternative growth substrate for hiPSCs. From the initial screening of homopolymers, it was found that pluripotency maintenance was not affected by the polymer substrate, however, the number of attached cells varies greatly per polymer (Figure 4.13).



**Figure 4.14:** Average cell count versus pluripotency maintenance after microarray screening. The use of different polymer chemistry resulted in a change of cell attachment and pluripotency maintenance. Data from Nasir *et al.* [59]

In total, 216 polymers were fabricated in triplicate onto separate Droplet Microarrays (Figure 4.14). After 7 days of drying in the vacuum oven (35 °C ; <50 mTorr), the array was washed in deionised H<sub>2</sub>O for 24 hours and analysed with ToF-SIMS. The array was thereafter incubated for 1 hour at 37 °C in 50 mL 1x Essential 8™ and 10 µM ROCK inhibitor Y-27632 in DMEM/F12. After removal of non-adsorbing components through three 10-second dip washes in deionised H<sub>2</sub>O, *in situ* digestion was performed for 4 hours at 37 °C with 0.05 sequencing grade trypsin (Section 3.4.4). The array was then dried in a vacuum oven (<50 mTorr) for five minutes to remove the DMSO. Next, the array was analysed via LESA-MS/MS.



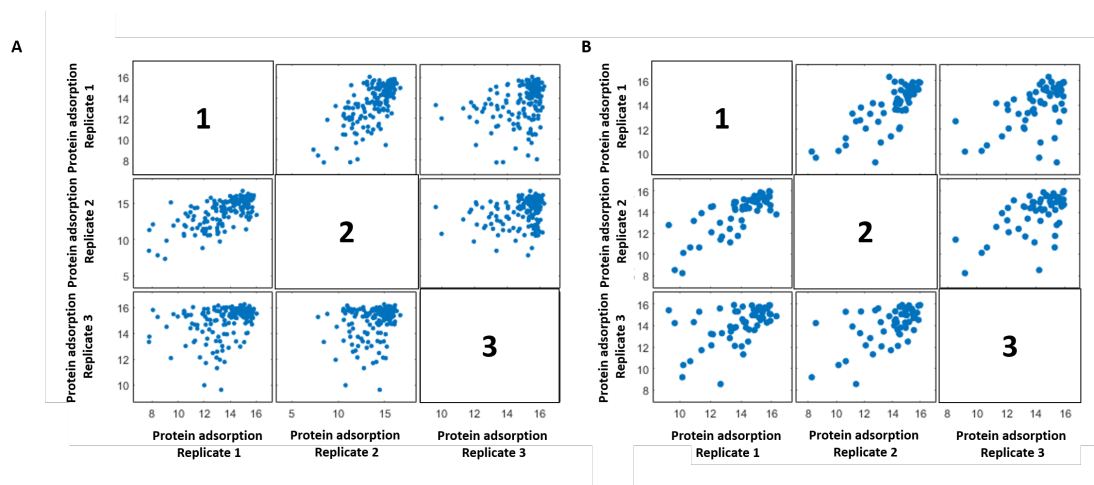
**Figure 4.15:** Sequence for printing selection monomer onto the Droplet Microarray (ø1.414 mm)

ToF-SIMS analysis showed that 180 out of the 216 polymers were printed successfully onto the Droplet Microarray. Whilst the solvent system and printing parameters were optimised for a range of monomer solutions, solutions can still be repelled by the superhydrophilic surface or dissolving the superhydrophilic surface and therefore preventing *in situ* polymerisation.

After applying Thompson's tau test on the cell data, 390 outliers were removed. Then, Thompson's tau test was also applied on the log-transformed quantile-normalised protein data after which both datasets were subjected to an abundance filter. The number of outliers per protein are shown in Table 4.5. The effect of normalisation and filtering is shown in Figure 4.15. It can be seen that correlation between replicates improved as a result of the data pre-processing steps.

**Table 4.5:** Number of outliers (Thompson tau) and removed polymers (abundance filter) per Essential 8<sup>TM</sup> protein after filtering

Protein	Outliers protein adsorption
Insulin	31
Transferrin	23
FGF-2	23
TGF-β1	2



**Figure 4.16:** Effect of data pre-processing steps on variability between replicates. (A) Pair-wise comparison of all log-transformed data for insulin adsorption. (B) Pair-wise comparison after quantile normalisation and outlier filtering using Thompson's tau test. Improvement in agreement between replicates is observed, i.e. data clean-up led to improved correlation between replicate (better agreement)

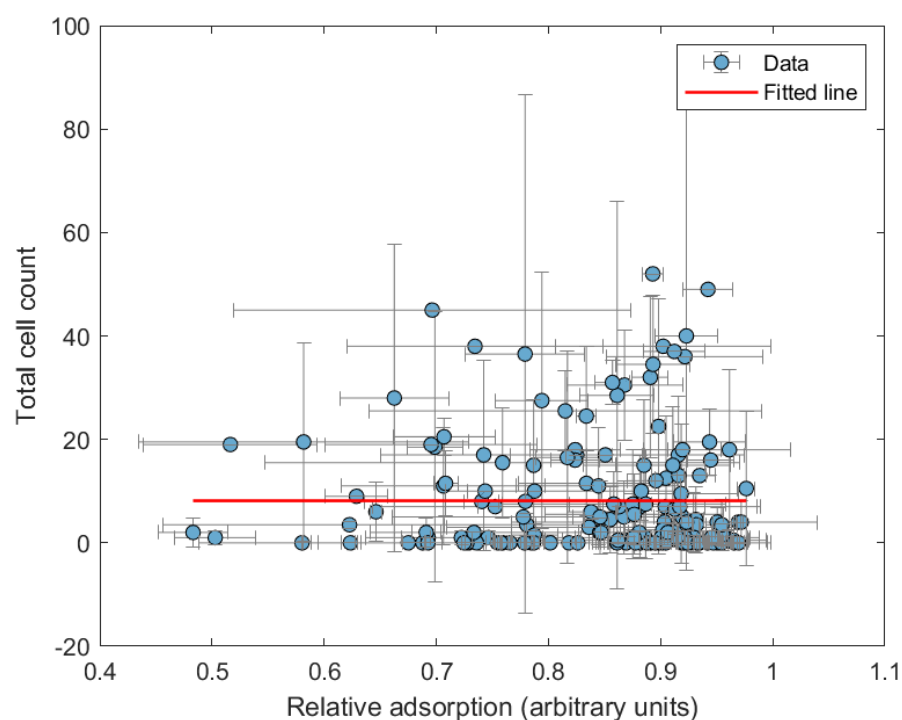
Filtered data were statistically compared to identify any difference in protein adsorption between polymer substrates. It was found that the data did not follow a normal distribution (Lilliefort's test:  $p < 0.001$ ), therefore, non-parametric comparison (Kruskal-Wallis) was performed on the protein adsorption data. For all Essential 8<sup>TM</sup> proteins, differential adsorption was observed between polymer chemistries (Table 4.6).

**Table 4.6:** Comparison of protein adsorption values for Essential 8<sup>TM</sup> proteins across different polymers

Protein	p-value
Insulin	< 0.001
Transferrin	< 0.001
FGF-2	< 0.001
TGF- $\beta$ 1	< 0.001

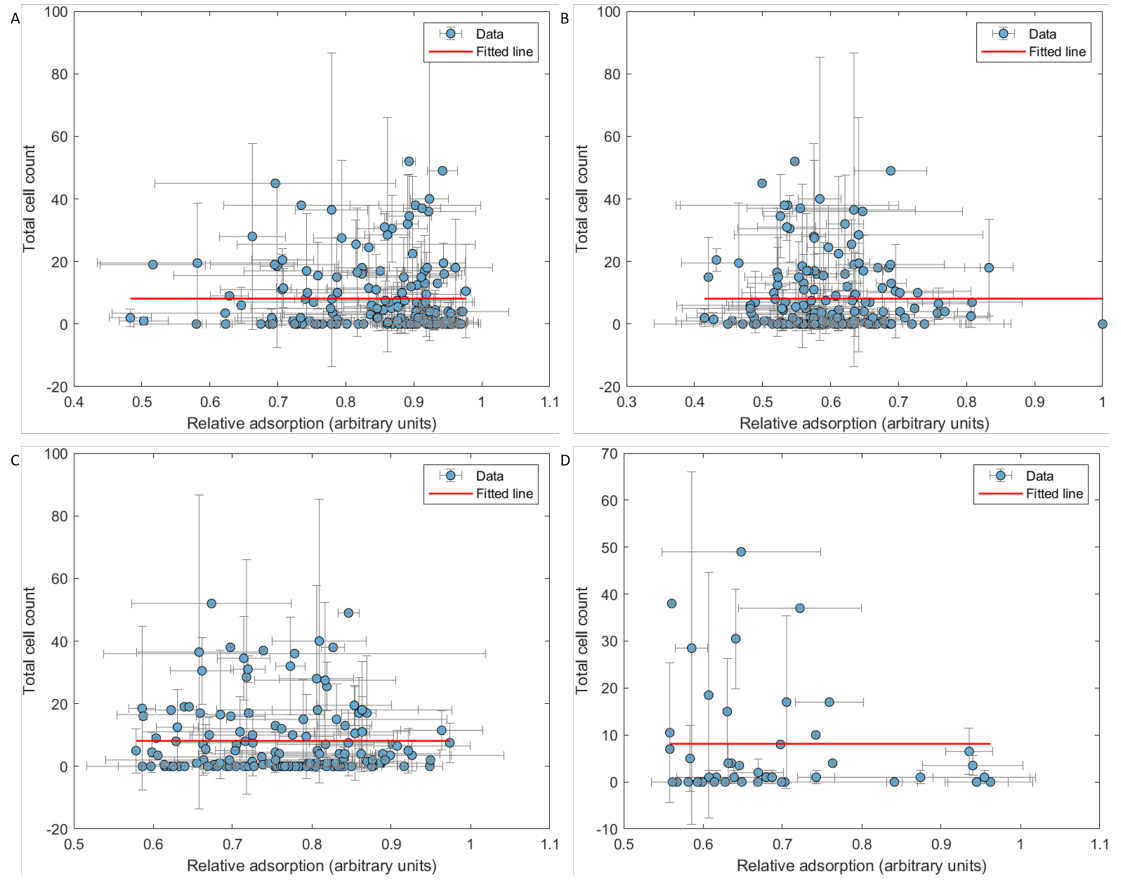
After, regression analysis was performed using the Curve Fitting Toolbox in MATLAB. Non-linear regression was performed using a sigmoid function. A sigmoid function was used as a sigmoid profile for fibronectin adsorption on different co-polymer

compositions was observed by Mei *et al.* [66]. Therefore, the same type of curve fitting ( $y = \frac{a}{b + e^{-x}}$ ) was used for this data. Initially, no correlations were found ( $r < 0.001$ ;  $p > 0.05$ ) between protein adsorption and total cell count. However, the large number of zero values in the cell data cause the trend line to not reflect any trend in the data (see Figure 4.16 for an example). Therefore, polymers which had a cell count of zero were removed from the data.



**Figure 4.17:** Non-linear regression results for all cell data and insulin adsorption. The high amount of zero values is masking the true trend in the data whilst showing high correlation

As can be derived from Figure 4.17A-D and the regression analysis, a sigmoid model does not accurately describe the relation between protein adsorption and cell attachment. However, visually it appears that the higher adsorption of TGF- $\beta$ 1 leads to a decrease in the attachment of pluripotent stem cells.



**Figure 4.18:** Scatter plots for cell attachment as a function of Essential 8<sup>TM</sup> protein adsorption. (A) Insulin, (B) transferrin, (C) FGF-2, and (D) TGF- $\beta$ 1 . All correlation coefficients were  $< 0.001$  and none of the regression models was significant ( $p > 0.05$ )

In the study of Mei *et al.* [66], a selection of co-polymers were incubated with  $0.03 \text{ mg} \cdot \text{mL}^{-1}$  fibronectin. For a number of co-polymers, a saturation effect was observed, i.e. a further increase in fibronectin adsorption did not lead to a further increase in cell attachment. For the remaining selection of co-polymers no obvious relation was seen between fibronectin adsorption and cell count.

Here, an opposite effect is observed, i.e. there is a decrease in cell attachment when the relative amount of TGF- $\beta$ 1 increases. Therefore, this data suggest that other factors than culture medium protein adsorption increase the cell attachment and pluripotency maintenance. These factors will be discussed in Section 4.4.5.

#### 4.4.4 INVESTIGATION OF STRUCTURAL RELATED PROTEIN ADSORPTION USING MOLECULAR DESCRIPTORS AND PLS REGRESSION

The relative protein adsorption data were and ToF-SIMS ion intensities were used to find any relation between protein adsorption and the chemical nature of the polymer. To do so, protein adsorption data were subjected to PLS regression. First, the number of components were selected to capture at least 90% of variance in the data. Then, a PLS model was built using that number of components and VIP scores were calculated. ToF-SIMS ions with a VIP score  $> 1.5$  were included in the final model. For the final PLS model, the optimal number of components (smallest mean square error of prediction) was estimated using a 5-fold cross validation strategy which was repeated 50 times. The predictivity of the PLS model (how well do ToF-SIMS ion intensities predict protein adsorption) was assessed by calculating the goodness-of-predictivity ( $Q^2$ ).

**Table 4.7:** Results of pre-processing protein adsorption data for building a predictive PLS model

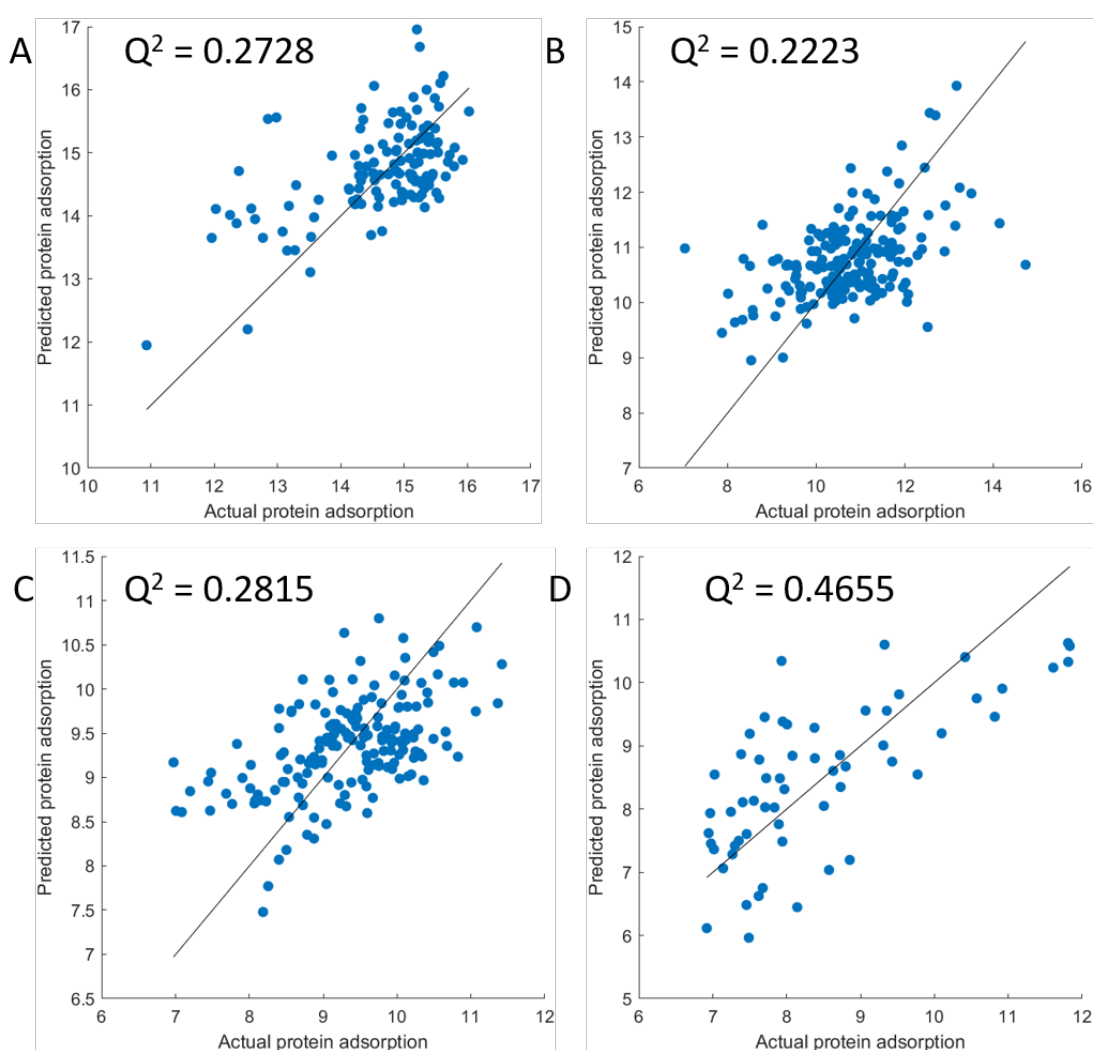
Protein	Number of components to explain $>90\%$ variance	Number of included ions final PLS model
Insulin	20	47
Transferrin	15	38
FGF-2	15	51
TGF- $\beta$ 1	10	53

For all proteins, a prediction model was obtained with a  $Q^2 > 0$ , indicating a certain degree of predictivity (Figure 4.18). The best prediction model (highest  $Q^2$ ) is obtained for TGF- $\beta$ 1 . These results show that protein adsorption can be related to the chemical structure of the polymeric substrate.

For a biologically relevant model a  $Q^2 > 0.4$  is required [208]. In order to assess whether all models have predictive value, PLS regression was performed by running permutations ( $n = 100$ ) [250]. In a permutation round, the protein adsorption data is randomly assigned to the ToF-SIMS ion intensities. PLS regression is then performed



using the same number of components and  $Q^2$  is then calculated through leave-one-out cross validation. The  $Q^2$  values of the original model and the random models are then statistically compared using a z-test. A relevant prediction PLS model should perform significantly ( $p < 0.05$ ) better than any random PLS model. The results of the permutation test are shown in Table 4.8. These results show that for each protein the PLS model predicts protein adsorption better than any random model. Therefore, all PLS models are relevant to predict protein adsorption from ToF-SIMS ion intensities.



**Figure 4.19:** Prediction of Essential 8<sup>TM</sup> protein adsorption using ToF-SIMS ion intensities and PLS regression. (A) insulin (B) transferrin (C) FGF-2 and (D) TGF- $\beta$ 1. The black line represents the identity ( $x = y$ )

**Table 4.8:** Results of permutation test for predictive PLS models

Protein	p-value
Insulin	< 0.001
Transferrin	< 0.001
FGF-2	< 0.001
TGF- $\beta$ 1	< 0.001

To investigate which structural features are important variables in the respective PLS models, ions with a  $m/z < 100$  were assessed. Assignment of putative identities was done using prior work from Hook & Scurr [254]. The list of important ions per protein are shown in Table 4.9.

**Table 4.9:** Important ions in PLS model for predicting Essential 8<sup>TM</sup> protein adsorption. Reported values are  $m/z$  values

Insulin	Transferrin	FGF-2	TGF- $\beta$ 1
22.99	24.13	24.13	16.09
24.13	28.02	27.02	28.02
40.01	77.95	46.98	32.02
46.98	86.99	53.99	40.01
60.41	89.01	63.95	41.99
78.00	97.70	65.98	42.05
83.81	97.98	79.94	44.99
93.00	80.01	59.20	
94.02	83.81	65.03	
94.24	95.95		
97.70	96.94		
97.98			

For insulin, ions related to hydrophobic structures are observed (e.g.  $m/z$  40.1:  $C_3H_4$ ,  $m/z$  78.00:  $C_6H_6$ ,  $m/z$  93.00:  $C_6H_5O$ ). This could be expected since insulin is a hydrophobic protein (GRAVY: 0.22) and would therefore prefer to adsorb onto hydrophobic

surfaces. Transferrin (GRAVY: -0.41), FGF-2 (GRAVY: -0.53) and TGF- $\beta$ 1 (GRAVY: -0.25) are more hydrophilic proteins. Therefore it is expected that those protein would preferably adsorb onto more hydrophobic polymers. The ions  $m/z$  27.02 ( $C_2H_3$  or CN) and  $m/z$  28.02 ( $C_2H_4$ ) can be derived from heteroatom-rich structures from certain (meth)acrylate polymers. Furthermore,  $m/z$  79.94 relates to  $SO_3$  which is also known to be a hydrophilic group. Altogether, this shows that the observed important ions for predicting protein adsorption can be put into context based on hydrophobic-hydrophobic or hydrophilic-hydrophilic interactions.

#### 4.4.5 POTENTIAL FACTORS AFFECTING CELL RESPONSE

One potential factor of increased cell attachment at maximum protein adsorption is the presence of ROCK inhibitor Y-27632 (Figure 4.19). The ROCK inhibitor was added for the first 24 hours in the Essential 8<sup>TM</sup>. Y-27632 is known to promote cell attachment by increasing the spreading of the cells as well as through cofilin activation [255]. This could indicate that there might be an effect of Y-27632 on the attachment to protein-coated polymers. However, further investigation is required to assess the role of Y-27632 ( $m/z$  248.1763  $[M+H]^+$ ) since it did not fall within the currently used  $m/z$  range ( $m/z$  400-900).

In addition, for further understanding of the cell-material response mechanism, incubation experiments should be conducted in the same format but in the presence of cells. Attached cells should then be carefully removed without compromising the deposited matrix and without release of intracellular components. So far, studies (e.g. [66, 67, 69, 256, 257]) have only focused on adsorption of medium components without the presence of cells. Abdallah *et al.* found laminin ( $\alpha$ ,  $\beta$ ,  $\gamma$ ) and nidogen from Matrigel<sup>TM</sup> to adsorb to poly(methyl methacrylate) [69]. Differences in deposited matrices on synthetic polymers could then provide more understanding on the difference cell response for different synthetic polymer substrates. This would require the development of a robust protocol for cell removal as well as further development of the current LESA-MS/MS analysis strategy to permit quantification of more complex matrices.

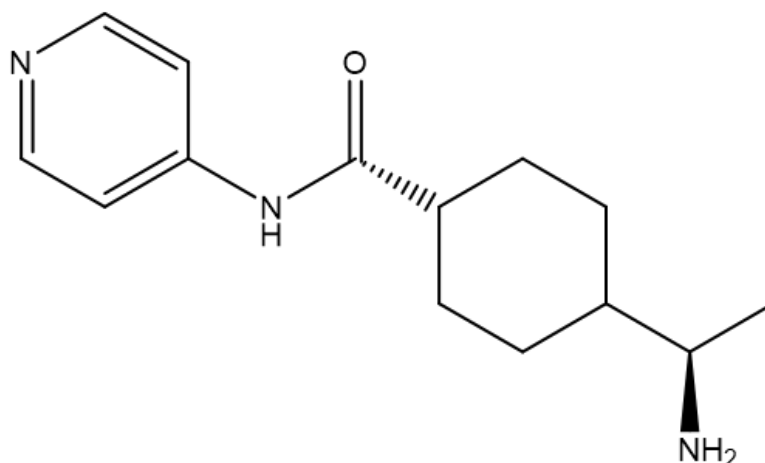


Figure 4.20: Chemical structure of ROCK inhibitor Y-27632

## 4.5 CONCLUSION

With the developed LESA-MS/MS method, quantitative analysis of protein adsorption on synthetic polymer surfaces was carried out successfully. From the differential adsorption profiles, the data suggest that the amount adsorbed protein does not correlate with cell attachment of pluripotent stem cells. Furthermore, difference in protein adsorption were used to identify molecular descriptors to suggest what molecular features induces/repel protein adsorption.

This method shows huge potential to further investigate monomer libraries to understand cell-material response and predict protein adsorption from molecular descriptors. This is vital information to select better performing structures to be used as growth substrates for hIPSCs. In addition, the methodology could also be extend to different cell lines and/or more complex media to improve upon the design of synthetic biomaterials.

Since a random scatter was observed for proteins with cell response, it suggests that the observed cell response cannot be solely attributed to adsorption of culture medium proteins onto the polymer substrates. Further study would require to look into a potential role of protein orientation/conformation or cell-induced posttranslational modifications in relation to hIPSC attachment.

# Quantitative screening of protein adsorption on scaled-up polymers using LESA-MS/MS analysis<sup>1</sup>

## 5.1 INTRODUCTION

**P**RE-ADSORPTION of medium-derived proteins present in cell culture media increases the cell response on synthetic polymer surfaces [67, 258]. High throughput experiments for discovery of a suitable fully-synthetic substrate for hPSC expansion in a minimal and fully-defined medium (Essential 8<sup>TM</sup> [259]) have shown a number of candidates to replace currently used biological feeder layers as Matrigel<sup>TM</sup> [59]. However, many potential hits derived from array screening experiments fail in the scale-up stage. During the scale-up stage, polymer coatings are created on a larger surface area (e.g. 96-well or 6-well plate) to test cell response under representative conditions.

In the previous chapter (Chapter 4), a high throughput strategy was developed for identification and quantification of proteins adsorbed onto synthetic polymers. Now, the aim is to transfer this methodology (LESA-MS/MS) for the analysis of scaled-up

---

<sup>1</sup>This chapter is part of the paper: A. Nasir *et al.*, *Adv. Healthc. Mater.* (2020), doi:10.1002/adhm.202001448.

polymers in order to replace labour-intensive and time-consuming LC-MS/MS analysis [67, 258]. Quantitative assessment of protein adsorption of selected polymers could potentially inform on the failure/success of scaled-up polymers. In this chapter, a selection of synthetic (co-)polymers in a 6-well plate format were subjected to quantitative LESA-MS/MS analysis to investigate the relation between protein adsorption and polymer performance regarding hPSC attachment and survival [59].

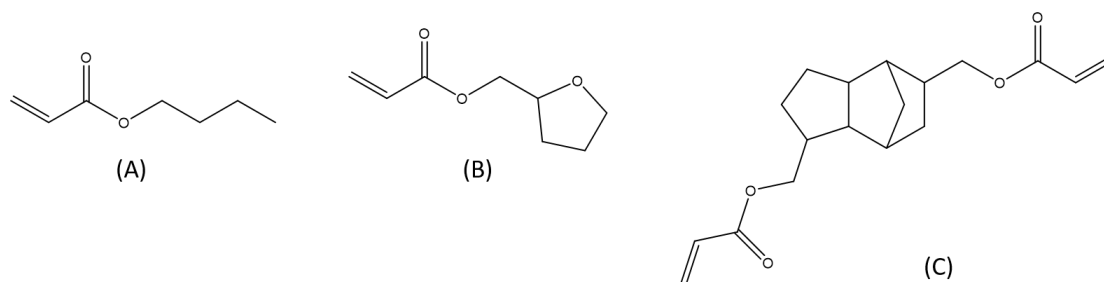
## 5.2 AIMS & OBJECTIVES

The aim of this chapter is to use the developed MS strategy to quantitatively study protein adsorption on a set of low, medium and high performing synthetic polymers for stem cell expansion.

## 5.3 MATERIALS & METHODS

### 5.3.1 REAGENTS & CHEMICALS

Butyl acrylate (BA;  $\geq 99\%$ ; Figure 5.1A), tetrahydrofurfuryl acrylate (THFuA; Figure 5.1B, 2-DMPA (99%), AmBic (BioUltra,  $\geq 99.0\%$ ), BSA ( $\geq 96\%$ ), horse heart myoglobin ( $\geq 90\%$ ), bovine heart cytochrome C ( $\geq 95\%$ ), and human *holo*-transferrin ( $\geq 97\%$ ) were purchased from Sigma-Aldrich (Gillingham, UK). Tricyclodecane-dimethanol diacrylate (TCDMDA; 95-100%; Figure 5.1C) was obtained from Polysciences GmbH (Hirschberg an der Bergstrasse, Germany). ACN and water (LC-MS grade; CHROMASOLV®) were bought from Riedel-de Haen. Permanox™ cell culture slides (Nunc®), Essential 8™ 50x supplement and Gibco DMEM/F12 were acquired from Thermo Fisher Scientific (Loughborough, UK). ROCK inhibitor Y-27632 hydrochloride (ROCKi) was obtained from STEMCELL™ Technologies (Vancouver, Canada). Sequencing grade trypsin was purchased from Promega (Southampton, UK)



**Figure 5.1:** Chemical structures of polymers. (A) BA, (B) THFuA, (C) TCDMDA

### 5.3.2 6-WELL PLATE FABRICATION

Polymer-coated 6-well plates were kindly provided by Dr Jordan Thorpe. Briefly, tissue culture plastic 6-well plates (Thermo Fisher, Hemel Hempstead, UK) were treated with oxygen plasma for 10 minutes at 100 W and  $2 \times 10^9$  mbar (Diener, Germany). After, monomer solutions (BA, THFuA, TCDMDA:BA 2:1 (v/v)) were mixed with 10% w/v 2-DMPA in IPA in a 9:1 v/v monomer-to-photoinitiator ratio. Solutions were degassed for 15 minutes at 30°C using a sonicator. Oxygen was removed from the solution vials by flushing with argon gas. Plasma-etched 6-well plates were filled with 125  $\mu$ L monomer-photoinitiator solution per well followed by 1 hour irradiation with UV light. To remove unreacted monomer, wells were rinsed three times with IPA, followed by three rinses with deionised water. Wells were filled with deionised water and incubated for 48 hours at 37°C. After 24 hours, wells were rinsed five times with deionised water and wells were filled with fresh deionised water. After 48 hours, the deionised water was discarded and polymer-coated well plates were stored at room temperature.

### 5.3.3 INCUBATION OF POLYMER-COATED WELL PLATES

Each (co-)polymer was tested in triplicate. The culture medium was prepared by mixing 49 mL of DMEM/F12 with 1 mL Essential 8<sup>TM</sup> 50x. A solution of 10  $\mu$ M ROCKi in deionised H<sub>2</sub>O were provided by Dr Jordan Thorpe and Dr Aishah Nasir. ROCKi solution was mixed with Essential 8<sup>TM</sup> 1x in a 1:1000 v/v ratio. Next, each well was

filled with 1.5 mL ROCKi/Essential 8<sup>TM</sup> 1x solution and incubated for 1 hour at 37. After, ROCKi/E8 1x solution was discarded and all wells were washed using an adapted protocol from Hammad *et al.* [67]. The washing cycle was carried three times and consisted of adding 1.5 mL fresh deionised H<sub>2</sub>O , reciprocal shaking for 10 minutes (Stuart reciprocal shaker; 100 spm), and discarding washing solution.

### 5.3.4 PROTEIN DIGESTION

#### 5.3.4.1 AMBIENT *in situ* DIGESTION

Sequencing grade trypsin (0.5 in 50 mM acetic acid) was diluted 10-fold with 100 mM AmBic. The working solution was dispensed onto pre-selected regions of interest. Incubation took place for 24 hours at room temperature after which the remaining trypsin solution was allowed to evaporate in a fume hood. Plates were stored at 4°C until further analysis

#### 5.3.4.2 MICROWAVE-ASSISTED *in situ* DIGESTION

Sequencing grade trypsin (0.5 in 50 mM acetic acid) was diluted 10-fold with 100 mM AmBic. The working solution was dispensed onto pre-selected regions of interest. Microwave-assisted *in situ* digestion was performed according to an adapted protocol of Ha *et al.* [260]. The sample was directly placed in a bowl with ice. Microwave-assisted *in situ* digestion was carried in a domestic microwave (Swan Compact 20701, Swan Housewares Ltd.; maximum power: 500 W) at medium power for 10 minutes. After, the trypsin solution was allowed to evaporate under ambient conditions.

### 5.3.5 LESA-MS/MS

Well plates were individually analysed using the TriVersa Nanomate (Advion Biosciences, Inc., Ithaca, NY) in 'LESA' mode. The stage temperature was set to 20. The extraction solvent consisted of 10% v/v ACN in 200 mM ammonium acetate [181]. In total, 5 µL extraction solvent was aspirated from the solvent reservoir from which 3.0

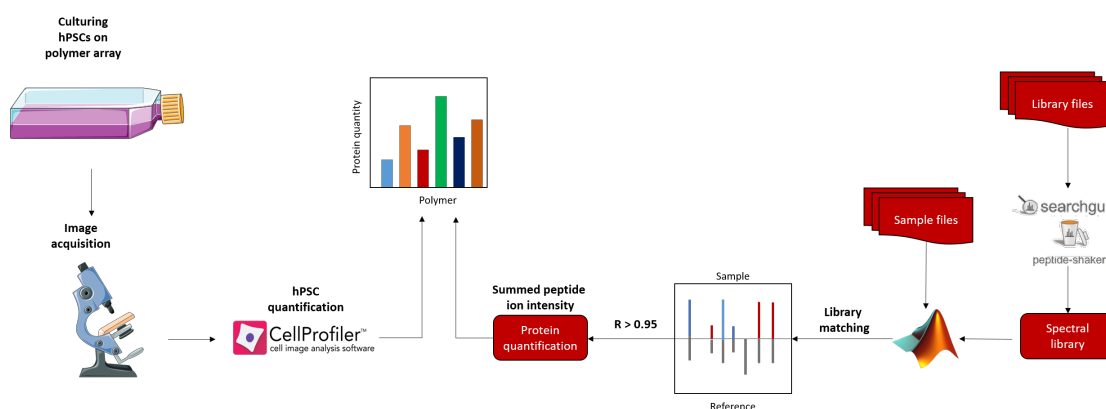


$\mu\text{L}$  was dispensed onto the digest for 10 seconds. Then, 3.5  $\mu\text{L}$  was aspirated back into the tip after which the sample was introduced into a Q Exactive plus (Thermo Scientific, San Jose, CA) via nanoelectrospray ionisation (ESI Chip<sup>TM</sup>, Advion Biosciences, Inc., Ithaca, NY) using 1.6 kV voltage and 0.4 psi gas pressure ( $\text{N}_2$ ). The mass spectrometer was operated in DDA mode. MS<sup>1</sup> spectra were acquired at a resolution of 140,000 (at  $m/z$  200) over the range  $m/z$  400-1200. The AGC target and ion injection time were respectively set to  $3 \times 10^6$  and 200 ms. MS/MS spectra were acquired at a resolution of 17,500 (at  $m/z$  200) with a maximum ion injection time of 250 ms and the AGC target set to  $1 \times 10^5$ . Ions were isolated within a 1  $m/z$  window and subsequently fragmented using a NCE of 27.

### 5.3.6 PROTEIN IDENTIFICATION & QUANTIFICATION

All processing was done using a custom MATLAB script ([https://github.com/jorismeurs/LESA\\_Proteomics](https://github.com/jorismeurs/LESA_Proteomics)). A spectral library was created from previously acquired MS/MS data of all Essential 8<sup>TM</sup> proteins. First, .RAW files were submitted to SearchGUI v3.3.20 [217] and spectrum matching was performed using the X!Tandem [104], MSGF+ [108] and OMSSA [103] algorithms. The following parameters were used for spectral matching: 10 ppm precursor mass tolerance, 0.02 Da fragment mass tolerance, fully tryptic peptides with a maximum of two miscleavages, and the FDR set to 1% for protein, peptide and PSM identification. A library file containing the spectra of the identified peptides was then generated in MATLAB and stored in .mat format. Analysis files were then converted to MASCOT Generic Format (.mgf) files using msconvert [199] with peak picking enabled. Sample MS/MS spectra were subsequently matched against the spectral library. Fragments tolerance was set to 0.02 Da and sample spectra with a cosine correlation greater than 0.95 were considered identified. The log-transformed averaged summed intensity of identified peptide ions were then retrieved and used for quantification. Intensity values were presented relative to the intensity of BA (lowest cell attachment) to investigate the correlation between protein adsorption and polymer performance regarding cell attachment and pluripotency maintenance. An overview

of the workflow is shown in Figure 5.2



**Figure 5.2:** Schematic workflow for performing quantitative LESA-MS/MS analysis for proteins adsorbed onto polymer-coated 6-well plates. Cells are seeded and cultured on the selection of polymers and subsequently stained with DAPI and OCT4 to quantify the total number of cells and number of pluripotent cells respectively. Cell quantification was performed using fluorescence microscopy and Cell Profiler [61] for automated quantification. In parallel, protein adsorption experiments were done on the same selection of polymers. A spectral library is first generated for the Essential 8™ proteins. The identified peptide ions were used to create a list of target peptides allowing ~ 1 minute analysis per sample. MS/MS spectra for the sample spectra were compared against the library. Then, protein quantification was done using the sum of the identified peptide ion intensities. To assess difference in protein adsorption between different polymer, ANOVA was used.

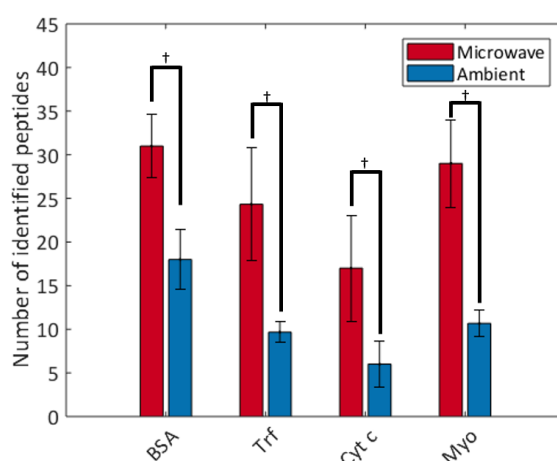
### 5.3.7 STATISTICAL ANALYSIS

Adsorption data were subjected to Grubb's test [261] to identify and remove any outliers. Missing values were thereafter imputed using the weighted average of  $k$  nearest neighbours ( $k = 11$ ) [249, 262]. ANOVA followed by Tukey's *post hoc* analysis was used for statistical comparison of protein adsorption. P-values less than 0.05 were considered significant.

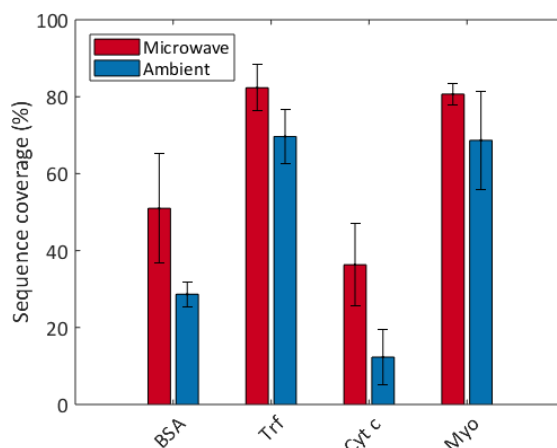
## 5.4 RESULTS & DISCUSSION

### 5.4.1 COMPARISON OF AMBIENT *in situ* DIGESTION AND MICROWAVE-ASSISTED *in situ* DIGESTION

Stock solutions for cytochrome C, transferrin, BSA and myoglobin (all  $1 \text{ mg} \cdot \text{mL}^{-1}$  in  $100 \text{ mM}$  AmBic) were prepared and mixed in a 1:1:1:1 (v/v/v/v) ratio. Protein mixtures were dispensed on Permanox<sup>TM</sup> cell culture slides either *in situ* digested under ambient conditions or through microwaving. Both digestion conditions were tested in triplicate. To assess the digestion efficiency, the number of peptides as well as the sequence coverage (summed peptide length divided by protein length) were assessed. For each protein, the number of identified proteins significantly increased when performing the digestion in a microwave (Figure 5.3). As a result of the identification of more peptides, the sequence coverage for each protein was higher when *in situ* digestion was performed in a microwave (Figure 5.4). A higher sequence coverage has been reported to be an indication of a more efficient digestion [79]. Furthermore, the variation in results was not found to be different for any of proteins between digestion conditions (*F*-test:  $p > 0.05$ ).



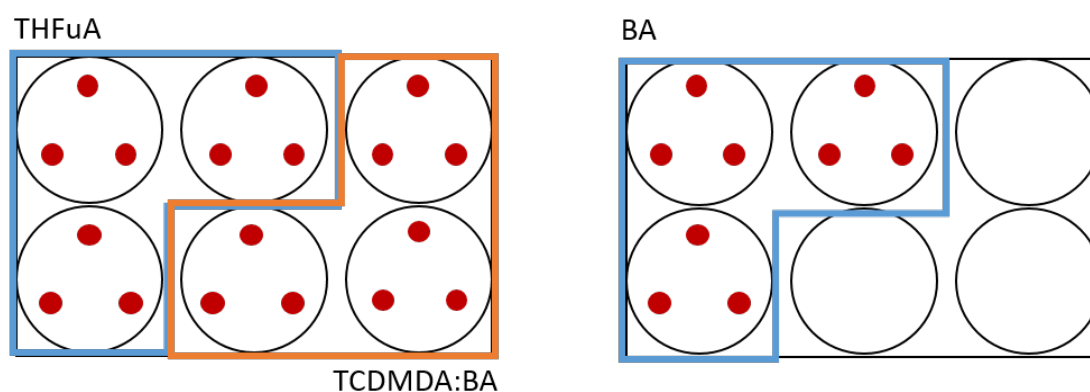
**Figure 5.3:** Comparison of number of identified peptides between ambient *in situ* digestion and microwave-assisted *in situ* digestion. †:  $p < 0.05$



**Figure 5.4:** Comparison of sequence coverage between ambient *in situ* digestion and microwave-assisted *in situ* digestion

#### 5.4.2 IDENTIFICATION OF ESSENTIAL 8™ PROTEINS FROM POLYMER-COATED WELL PLATES

Polymer-coated 6-well plates were incubated with a ROCKi/Essential 8™ 1x solution (1:1000 (v/v)) for 1 hour, plates were washed, followed by extraction and digestion of the adsorbed proteins. Well plates were designed to have two polymers per plate with three replicates. Each replicate was analysed in triplicate using LESA-DDA-MS/MS analysis (Figure 5.5).



**Figure 5.5:** Experimental design for quantitative proteomics on a selection of synthetic substrates

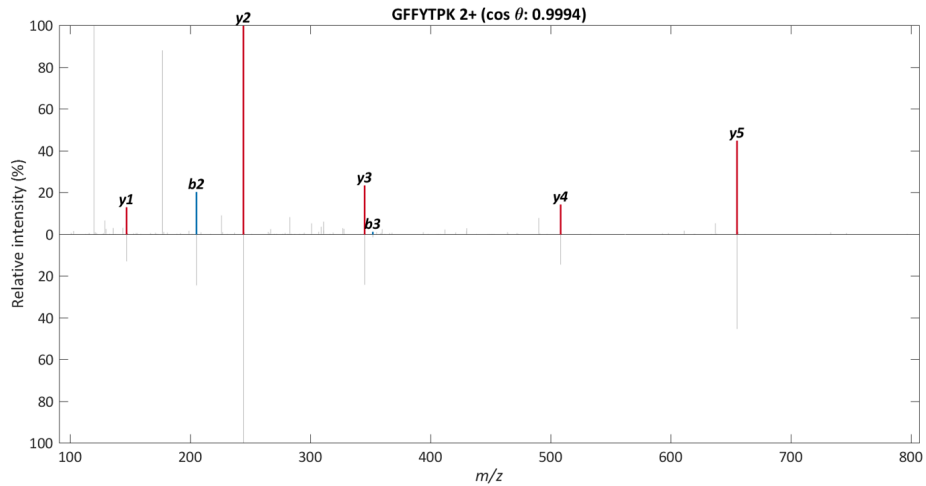
A spectral library with previously acquired DDA files on Essential 8™ was build

using LESAProteomics and used for identification of spectra acquired from the digests. Library identification has shown to produce more robust results [214]. Like in Chapter 4, MS<sup>1</sup> peptide ion intensities were summed, averaged and log-transformed per protein. Identification of DDA MS/MS spectra was done using the normalised dot product (cosine distance; Equation 5.1) between the reference (library; Appendix A) and sample spectrum. For identification, the following criteria were set:

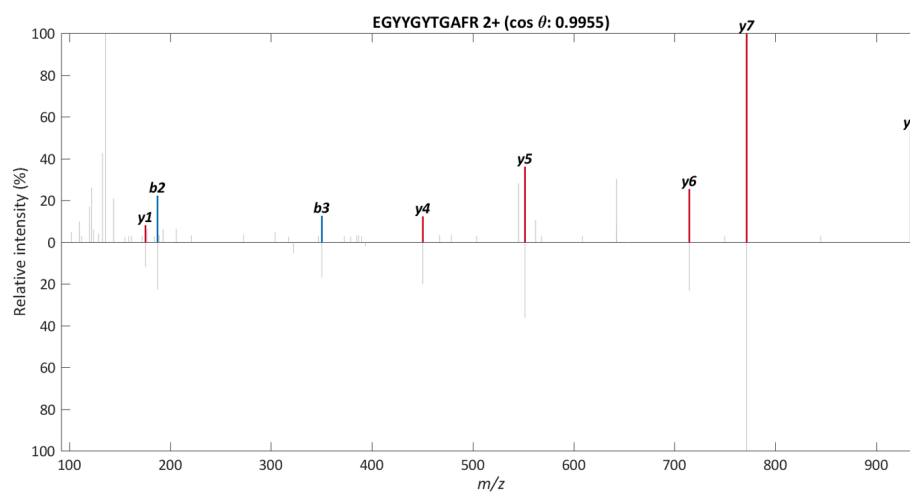
- $\cos \theta > 0.95$
- $\geq 5$  fragment ions could be annotated (mass tolerance window: 0.02 Da)

$$\cos \theta = \frac{\sum_{i=1}^n I_{sample} \times I_{reference}}{\sqrt{\sum_{i=1}^n I_{sample}^2 \sum_{i=1}^n I_{reference}^2}} \quad (5.1)$$

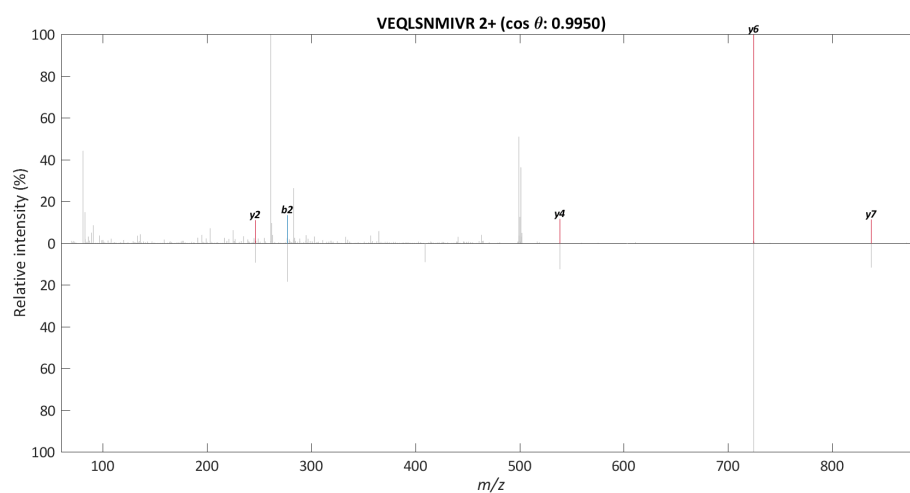
Insulin (Figure 5.6), transferrin (Figure 5.7), TGF- $\beta$ 1 (Figure 5.8), and FGF-2 (Figure 5.9) could all be identified through library identification. This shows a huge potential to perform future analysis on incubated scaled-up polymer format to examine cell-material interactions.



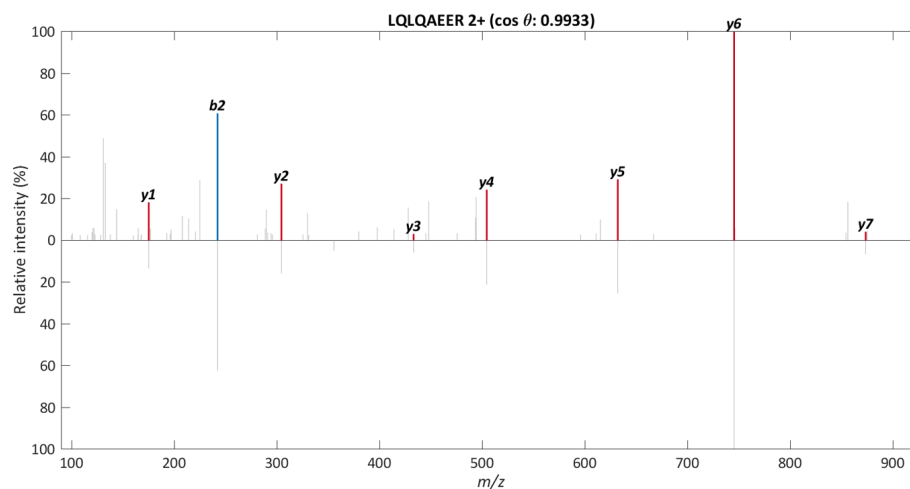
**Figure 5.6:** Identified tryptic peptide through library search for insulin



**Figure 5.7:** Identified tryptic peptide through library search for transferrin

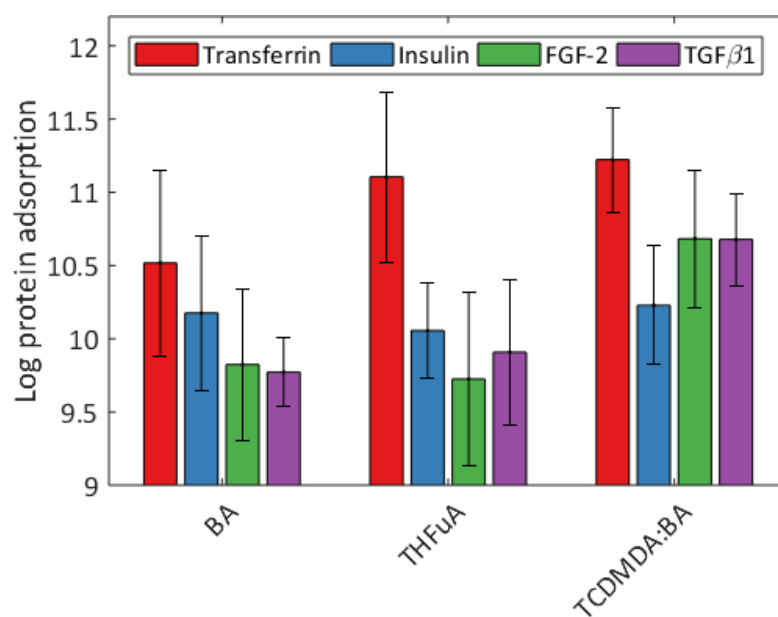


**Figure 5.8:** Identified tryptic peptide through library search for TGF- $\beta$ 1



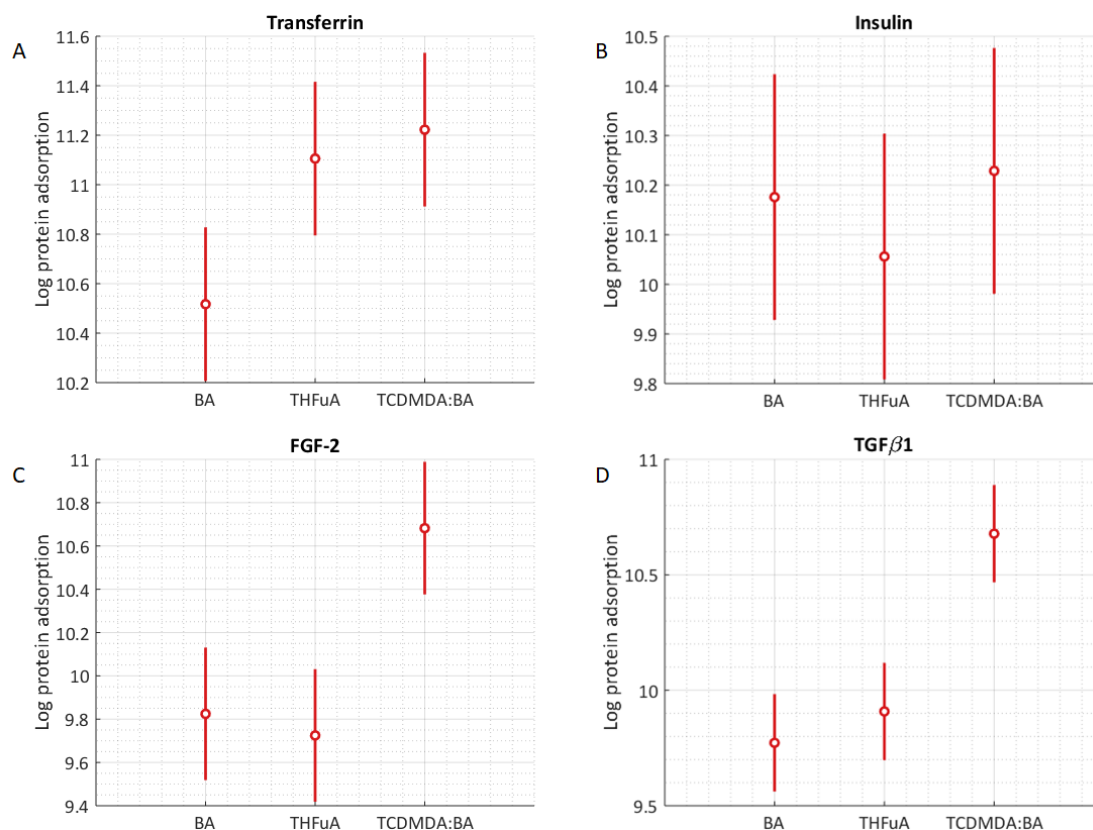
**Figure 5.9:** Identified tryptic peptide through library search for FGF-2

For all identified peptides per protein the intensity was retrieved from the averaged MS<sup>1</sup> spectra. For each protein, peptide intensities were averaged and log-transformed (Figure 5.10). Transferrin adsorption was higher on THFuA and TCDMDA:BA, but ANOVA followed by Tukey's *post hoc* analysis only revealed a significant higher adsorption of transferrin on TCDMDA:BA compared to BA (Figure 5.11A). For insulin, no difference in adsorption was observed between the different substrates (Figure 5.11). Additionally, it was observed that both FGF-2 and TGF- $\beta$ 1 adsorbed at higher quantities to TCDMDA:BA compared to THFuA and BA.



**Figure 5.10:** Comparison of protein adsorption for polymer-coated 6-well plates incubated with Essential 8™ medium. Data presented as mean values  $\pm$  standard deviation





**Figure 5.11:** Statistical comparison (ANOVA + Tukey's *post hoc* analysis) of protein adsorption on different synthetic polymer substrates. (A) Transferrin (B) Insulin (C) FGF-2 and (D) TGF-β1 . Non-overlapping intervals are significantly different

THFuA was the polymer with the second highest cell response in the screening of homopolymers, whilst TCDMDA:BA came forward as potential growth substrate for hIPSCs based on high cell response, long term (>72 h) cell culture survival and the ability to differentiate hIPSCs into endoderm, mesoderm and ectoderm lineage cells [59]. Interestingly, for TCDMDA:BA a significantly higher adsorption was observed of FGF-2 and TGF-β1 compared to BA and THFuA. Higher hIPSC attachment and long-term survival (>72 h) on TCDMDA:BA can be explained by a higher degree of TGF-β1 adsorption. TGF-β1 can bind integrins, which are important glycoproteins for cells to be able to bind to extracellular matrix proteins [263, 264].

For transferrin, higher adsorption was observed for polymers with higher cell response (THFuA and TCDMDA:BA), though, no difference was observed between THFuA

and TCDMDA:BA. This seems to be in concordance with the observation in Section 4.4.3 that transferrin adsorption reaches saturation after which potentially currently unknown other factors influence the cell attachment and survival.

Insulin did not differ in protein adsorption between the three synthetic polymer substrates. Both insulin and transferrin are important factors for cell attachment and survival [259]. Based on the observation in this experiment, all Essential 8<sup>TM</sup> proteins will adsorb to a polymeric substrate regardless of the cell response. Transferrin seems to be an initial important factor for cell attachment until the substrate is saturated whereafter the growth factor proteins FGF-2 and TGF- $\beta$ 1 start to play an important role. However, in order to investigate the role of the individual proteins, a study should be performed on pre-coated polymer substrates with individual Essential 8<sup>TM</sup> proteins as well as combinations of Essential 8<sup>TM</sup> proteins to discover which protein (combinations) thrive cell response, cell culture survival and trilineage differentiation. Furthermore, despite a number of studies using LESA-MS for MSI [171, 171, 180, 182, 265, 266], it is not suitable for imaging at micrometer resolution in order to study the spatial adsorption profile of Essential 8<sup>TM</sup> proteins across the polymer substrate. Recent developments in 3D OrbiSIMS might be a suitable approach, however, protein analysis using 3D OrbiSIMS is relatively new and needs to be further investigated [267]. MALDI MSI could be an alternative approach to acquire chemical images for *in situ* trypsin digested Essential 8<sup>TM</sup> proteins [268–271].

## 5.5 CONCLUSION

The developed LESA-MS/MS strategy was extended from array screening to the analysis of scaled-up polymers in a 6-well plate format. Through microwave-assisted *in situ* digestion, the digestion time was reduced from 4 hours to 10 minutes. Furthermore, using the TriVersa Nanomate to directly sample the *in situ* digested proteins reduced the analysis time to  $\sim 1$  minute.

Differences in protein adsorption were found between a low cell response substrate

## CHAPTER 5

(BA) and high cell response substrates (THFuA and TCDMDA:BA). These observations revealed valuable information towards understanding of long-term cell culture survival and trilineage differentiation. Therefore, this method shows potential to investigate medium-derived proteins for other cell lines and their adsorption to synthetic polymer substrates and provide with valuable information towards development of suitable synthetic growth substrates for individual cell types.

# Sequential 3D OrbiSIMS and LESA-MS/MS-based metabolomics for prediction of brain tumor relapse from sample-limited primary tissue archives <sup>1</sup>

## 6.1 INTRODUCTION

CENTRAL nervous system (CNS) paediatric tumours are the most prevalent type of solid cancer diagnosed in children and the leading cause of mortality among all cancers in children [272]. From the clinical and biological perspective, intracranial paediatric ependymomas remain enigmatic and challenging tumours to treat. Overall, the prognosis is poor with over 50% of tumours relapsing and less than 50% of children surviving this disease (5-year overall survival is 25% for patients who have relapsed) [273–275]. Though widely considered a ‘surgical disease’, a significant proportion of

<sup>1</sup>This chapter has been fully published: Meurs, J.; Scurr, D. J.; Lourdasamy, A.; Storer, L. C. D.; Grundy, R. G.; Alexander, M. R.; Rahman, R.; Kim, D.-H. Sequential 3D OrbiSIMS and LESA-MS/MS-Based Metabolomics for Prediction of Brain Tumor Relapse from Sample-Limited Primary Tissue Archives. *Anal. Chem.* **2021**, 93 (18), 6947–6954, doi 10.1021/acs.analchem.0c05087

patients experience relapse, even following complete surgical resection of the tumour. The identification of biological correlates of disease progression and patient-tailored therapeutic targets therefore remains a significant challenge in this disease. Understanding the biochemical nature of tumour development is of vital importance for the development of the next generation of treatments [273]. In a disease state, the human metabolome is affected by several factors and therefore provides an excellent source of information to investigate disease-related alterations in metabolism [276]. To do so, an untargeted metabolomics approach can be used to study molecular changes within and between tissue samples of different phenotypes [277, 278]. State-of-the-art metabolomics techniques allow the detection of hundreds to thousands of metabolites in a biological sample [279], where untargeted metabolomics of cancer tissue is undertaken using liquid chromatography-mass spectrometry (LC-MS) and gas chromatography-MS (GC-MS) [280, 281]. Chromatography-based strategies allow the identification of a vast number of metabolites; however, these techniques require 20-50 mg of tissue for metabolomics analysis.

Preserving tumour regions of interest is commonly achieved using the tissue microarray (TMA) format. The neuropathologist identifies and cuts out key regions in the whole tissue section which are stored as a separate formalin-fixed paraffin-embedded (FFPE) block. The TMA platform allows small amounts of tissue to be used for transcriptomic and histological analysis [282–284], though for diagnosis the entirety of the tumour needs to be reviewed. Since the TMA tissue sections are small (diameter <1 mm; thickness: 4  $\mu$ m), sensitive analytical techniques are required for metabolomic studies to detect low-abundance metabolites. Despite the vast amount of available TMA libraries, metabolite profiling of tumour TMAs has been an unexplored territory due to incompatibility of LC-MS or GC-MS analysis. To perform MS analysis on TMAs, a sensitive technique is required to directly obtain a wide range of metabolites from small tissue sections.

A few studies have shown the potential of liquid extraction surface analysis-MS (LESA-MS) for untargeted metabolomics across a range of sample types. With LESA-

MS, liquid microjunction-based extraction can be performed on a flat sample surface to obtain the analytes of interest which are directly injected into the mass spectrometer [153]. Hall *et al.* found significantly changed profiles of lipids in non-alcoholic fatty liver disease (NAFLD) tissue using LESA-MS [198]. This allowed discrimination between different stages of steatosis. Ellis *et al.* used LESA-MS for the analysis of single-cell arrays and could distinguish cell types based on lipid profiles showing the capability of performing single-cell metabolomics with LESA [193]. LESA-MS was also successfully used to identify metabolic changes in small volumes of urine samples from an intervention study [211]. Basu *et al.* used LESA-MS for direct metabolite profiling for several different breast cancer cell lines [285]. The capability of LESA was shown to allow direct analysis of adherent cells with minimal sample preparation. Collectively, these studies have shown that from a limited amount of sample, metabolic changes could be measured accurately.

The recent development of Orbitrap secondary ion mass spectrometry (3D OrbiSIMS) revealed new possibilities for metabolic profiling due to its capability of high mass accuracy and mass resolving power ( $>240,000$  at  $m/z$  200) at subcellular spatial resolution [41]. In that study, it was shown that 3D OrbiSIMS can be used for 2D and 3D imaging of neurotransmitters, *in situ* identification of lipid species using MS/MS, and performing metabolomics profiling of single cells [41]. One unmet scientific challenge for brain tumour research is the capability to perform metabolomics analysis on archived TMAs to understand tumor development and find potential targets for therapies [279, 286].

3D OrbiSIMS and LESA-MS/MS require only minimal sample preparation, analysis can be performed directly on the tissue sample, and both instruments can acquire data in an automated manner using the TMA as sample platform. These MS techniques can therefore circumvent the need for tissue homogenisation, allowing the tissue to remain architecturally intact and available for subsequent study. The data achievable using this approach will address current challenges in cancer metabolomics, as detection of low abundance (highly polar) oncometabolites to study important metabolic pathways may enable the development of novel prognostic and treatment strategies [280, 287].

To date, no disease studies have thus far reported the use of combined 3D OrbiSIMS and LESA-MS/MS for untargeted metabolite profiling on TMAs. Combining MS techniques, in which ions are generated via different mechanisms, will allow acquisition of complementary metabolomic datasets from the same set of samples. Combination of the individual datasets on existing TMA archives could therefore provide a vast amount of clinically valuable information. Here, we perform 3D OrbiSIMS and LESA-MS/MS analysis of FFPE pediatric ependymoma TMAs as an exemplar demonstration of the ability to perform untargeted surface metabolomics and obtain clinically relevant data.

## 6.2 AIMS & OBJECTIVES

The aim of this chapter is to develop a surface mass spectrometry strategy for the analysis of metabolites in FFPE tissue tumour microarrays using 3D OrbiSIMS and LESA-MS/MS.

## 6.3 MATERIALS & METHODS

### 6.3.1 REAGENTS & CHEMICALS

MeOH and H<sub>2</sub>O (LC-MS grade; CHROMASOLV) were purchased at Riedel-de Haen (Seelze, Germany). FA (LC-MS grade; Optima™) was bought from Fisher Scientific (Loughborough, UK). Xylene (mixture of isomers; 98.5%, AnalaR NORMAPUR®) was acquired from VWR (Leicestershire, UK).

### 6.3.2 EXPERIMENTAL DESIGN

Hematoxylin and eosin-stained sections from FFPE paediatric ependymoma were examined by a neuropathologist at Nottingham University Hospital, and three representative areas were marked on the slides. Using a Raymond Lamb tissue micro-arrayer, 1 mm cores were punched from the marked areas of the donor blocks and placed into

recipient paraffin blocks to generate a tissue microarray (TMA). 4- $\mu$ m sections were cut from each block for use in further experiments. The analyzed TMA blocks consisted of patients which experienced tumour relapse ( $N = 5$ ;  $n = 3$ ) and patients without relapse ( $N = 2$ ;  $n = 3$ ) (Table 6.1).

**Table 6.1:** Demographic information for patients included in the analysis ( $N = 7$ ;  $n = 3$ )

Parameter	Value
Age at diagnosis	$55.3 \pm 39.9$ months
Gender	Female ( $N = 5$ ); Male ( $N = 2$ )
Recurrence	Yes ( $N = 5$ ); No ( $N = 2$ )
Tumour location	Posterior fossa ( $N = 6$ ); Supratentorial ( $N = 1$ )
WHO tumour grade	II ( $N = 4$ ); III ( $N = 3$ )
Overall survival	$60.1 \pm 35.1$ months
Status	Alive ( $N = 4$ ); Death ( $N = 3$ )

### 6.3.3 SAMPLE PREPARATION FOR MS ANALYSIS

Deparaffinisation of FFPE paediatric ependymoma TMAs was achieved using an adapted protocol from Ly *et al.* [288]. FFPE TMAs were first washed twice for 1 minute in a xylene bath. Residual xylene was removed and the array was allowed to dry in a fume hood for at least one hour, prior to storage at room temperature until analysis.

### 6.3.4 3D ORBISIMS ANALYSIS

The TMA was placed in a hybrid TOF.SIMS 5 (IONTOF GmbH, Münster, DE) instrument coupled to a Q Exactive HF (Thermo Scientific, San Jose, CA) mass spectrometer. Ions were sputtered from the surface using a 20 keV  $\text{Ar}_{3000}^+$  gas cluster ion beam (GCIB). The field of view (FoV) was set to 500  $\mu\text{m} \times 500 \mu\text{m}$ . Spectra were acquired at a resolution of 20  $\mu\text{m}$  in random raster mode. The Orbitrap was operated in Full-MS mode. The resolution was set to 240,000 at  $m/z$  200 and the AGC target was set to



$1 \times 10^6$  with a maximum ion injection of 511 ms. Data were acquired in the scan range  $m/z$  75-1125 for both positive and negative polarity.

### 6.3.5 LESA-MS/MS ANALYSIS

The TMA was placed on a universal plate holder (Advion Biosciences, Ithaca, NY) and scanned with an Epson V330 scanner. The tissue sample location was selected in LESA Points (Advion Biosciences, Ithaca, NY). Liquid extraction surface analysis-tandem mass spectrometry (LESA-MS/MS) was carried out using a TriVersa Nano-mate (Advion Biosciences, Ithaca, NY) coupled to a Q Exactive plus Orbitrap mass spectrometer (Thermo Scientific, San Jose, CA). Extraction of metabolites from tissue samples was conducted with a mixture of 80% v/v MeOH and 20% v/v H<sub>2</sub>O to which FA was added (end concentration 1% v/v). Brain tissue was sampled using the contact LESA approach [155] in which the solvent tip is brought into contact with the sample to minimize the solvent spread. During contact, 1.5  $\mu$ L solvent (total volume: 3  $\mu$ L) was dispensed on the tissue and after 15 seconds, 2.0  $\mu$ L was aspirated back into the tip. The extract was introduced into the mass spectrometer via chip-based nanoelectrospray ionisation (ESI Chip<sup>TM</sup>, Advion Biosciences, Ithaca, NY) at 1.4 kV and 0.3 psi gas pressure [211]. The mass spectrometer was operated in Full-MS/dd-MS<sup>2</sup> mode. MS<sup>1</sup> spectra were acquired in the scan range of  $m/z$  70-1050. The resolution was set to 140,000 at  $m/z$  200 and the AGC target was set to  $3 \times 10^6$  with a maximum ion injection time of 200 ms. Data-dependent MS/MS spectra were acquired at a resolution of 17,500 at  $m/z$  200. The AGC target for MS<sup>2</sup> scans was set to  $1 \times 10^5$  with a maximum ion injection time of 50 ms. The top 20 most intense ions were isolated within a 1  $m/z$  window for fragmentation. Dynamic exclusion was applied for 120 seconds per polarity. Fragmentation was carried using higher-energy collisional dissociation (HCD) using a stepped collision energy of respectively 10, 25 and 40 eV. All tissue sections were analyzed once. MS data were acquired for 2 minutes per polarity.

### 6.3.6 ION SELECTION FOR MULTIVARIATE ANALYSIS

LESA-MS spectra were averaged in XCalibur (Thermo Scientific, San Jose, CA) and exported in .RAW format. 3D OrbiSIMS spectra were exported from SurfaceLab (ION-TOF GmbH, Münster, Germany) as .TXT files. Mass spectrometry data were further processed using MATLAB (R2017a, The MathWorks, Inc., Natick, MA; <https://github.com/jorismeurs/LESAMS>). For LESA-MS data, files were converted to .mzXML using ProteoWizard 3.0.1908 [199]. Peaks were picked from the spectra using the `mspeaks` function from MATLAB's Bioinformatics Toolbox. After, peaks were aligned within 5 ppm  $m/z$  tolerance [198, 211]. Features with more than 20% missing values across all samples were removed [204]. Remaining missing values were imputed using  $k$ -nearest neighbours ( $knn$ ) imputation. The value of  $k$  was set to 10 [289].

### 6.3.7 MULTIVARIATE ANALYSIS AND PATHWAY ANALYSIS

Ion intensity matrices were TIC normalized, log-transformed, and fused using a low-level strategy [290]. Low-level fusion means that the obtained ion intensity matrices per method are concatenated and subjected to multivariate analysis as one dataset (Figure 6.1F). Data were divided into a *no relapse* for patients who did not experience relapse and *eventual relapse* for patients in who ependymoma recurred at least once after surgical removal. Fused data were subjected to partial least squares - discriminant analysis (PLS-DA) followed by leave-one-out cross-validation to calculate the performance of the model to distinguish between *no relapse* and *eventual relapse* samples. The PLS-DA model was further validated using permutation testing [250]. Ions with a VIP score greater than 1.5 were considered discriminative between groups and included for univariate statistical analysis. Discriminative ions were subjected to Student's  $t$ -test followed by a permutation test to estimate the FDR. Significant ions were annotated using the Human Metabolome Database (3 ppm mass tolerance) [291]. Annotated ions were submitted for pathway analysis in MetExplore [292, 293].

### 6.3.8 CHEMICAL HETEROGENEITY IN 3D ORBISIMS IMAGES

Identified metabolic classifiers for relapse were used to study intra-tumour heterogeneity in the 3D OrbiSIMS images. IONTOF .ITAX files were converted to .imzML and parsed into MATLAB using the imzML Converter [294] and an in-house developed MATLAB class ([http://github.com/jorismeurs/SIMS\\_imzML](http://github.com/jorismeurs/SIMS_imzML)). The dynamic range (difference between maximum and minimum ion intensity) was used as a measure for heterogeneity [295].

### 6.3.9 STATISTICAL ANALYSIS

Distribution of the data was assessed using Lilliefort's test. Univariate comparison between ependymoma subgroups was done using either Student's *t*-test or Mann-Whitney test for normal distributed and non-normal distributed data, respectively. A *p*-value less than 0.05 was considered to be significant. All statistical calculations were performed in MATLAB.

### 6.3.10 GENE EXPRESSION ANALYSIS

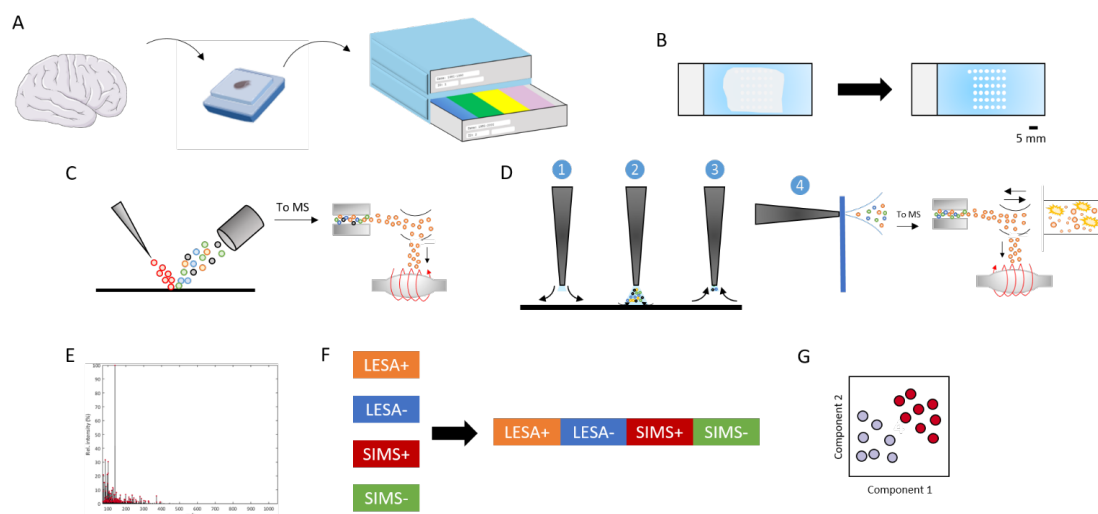
Two independent gene expression datasets from previously published studies were used for this analysis [296, 297]. The first cohort comprised of 65 tumour samples taken at the diagnosis as well as at relapse from 45 paediatric ependymoma patients. Transcriptomic microarray profiles of tumour samples were generated using Affymetrix HG-U133 Plus 2 GeneChip microarrays (Affymetrix, Santa Clara, CA). Expression intensity values were calculated at probeset level from microarray data CEL files using the robust multi-array average (RMA) method. Probesets that are 'absent' (present/-absent call using MAS5) in all samples were filtered out from the analysis. Expression values were mapped from probeset to unique gene and the probeset with the highest mean expression value was selected when multiple probesets were mapped to the same gene. The second cohort comprised of 54 tumour samples (both primary and relapse) from 17 paediatric ependymoma patients and profiled with whole-human-

genome 44K oligonucleotide microarrays (Dual Colour 44K microarray, Agilent Technologies). Raw intensities were background corrected and normalised with a Lowess normalisation method for within-array and quantile normalisation for between-array normalisation [298]. The normalised expression values were mapped to unique gene and the probe with the highest mean expression value was selected when multiple probes were mapped to the same gene. The normalised gene expression data from two studies were combined using ComBat, a robust empirical Bayes regression method [299]. The differential expression analysis between the primary and recurrent tumours with the ten candidate genes was performed with the moderated *t*-test and the gene set analysis was performed with the Global test [300].

## **6.4 RESULTS & DISCUSSION**

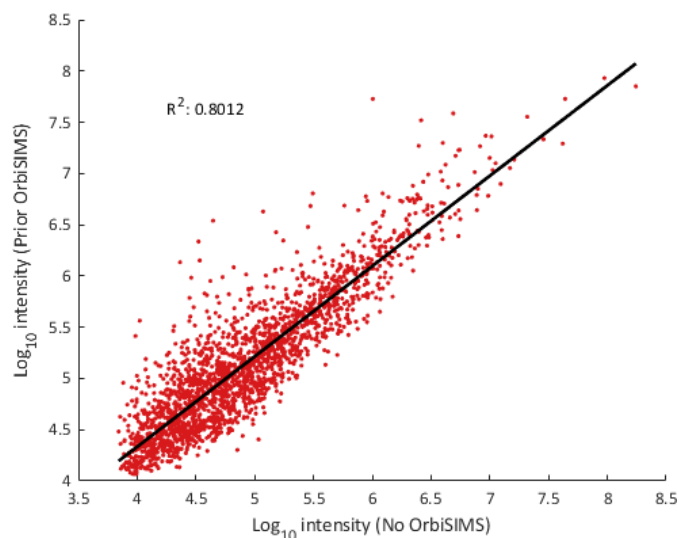
### **6.4.1 SERIAL 3D ORBISIMS AND LESA-MS/MS FOR IN SITU METABOLITE PROFILING**

The deparaffinised TMAs were first analysed with 3D OrbiSIMS followed by LESA-MS/MS and subsequent data preprocessing and multivariate analysis in MATLAB (Figure 6.1).



**Figure 6.1:** Sequential mass spectrometry analysis of pediatric ependymoma tissue microarrays. (A) The tumour tissue was removed, the tumour area was marked and then paraffin-embedded for long term storage. (B) For MS analysis, a TMA block from the archive was sectioned and mounted onto a glass substrate followed by a xylene wash to remove the paraffin. (C) Paraffin-free samples were then analysed by 3D OrbiSIMS followed by (D) LESA-MS/MS. (E) Ions were selected from the mass spectra and aligned. (F) All matrices with ion intensities were then combined (low-level data fusion). (G) Subsequently, data were subjected to partial-least squares-discriminant analysis (PLS-DA) to identify discriminative features in tumour recurrence. (H) Molecular formulae were assigned to the significant ions using the Human Metabolome Data-base. (I) Ions with a putative ID were then submitted to MetExplore for metabolic pathway analysis to identify affected pathways and corresponding genes.

The 3D OrbiSIMS was introduced to enable the imaging of metabolites within the individual TMA sections and to obtain the *in situ* spatial metabolite profile of FFPE ependymoma tissue. During SIMS analysis, the sample is slightly etched by the primary ion beam (20 keV  $\text{Ar}_{3000}^+$ ), although the amount of sample consumed by SIMS is limited when argon clusters are used [35]. Prior analysis by 3D OrbiSIMS did not lead to a decrease in either the number of detected ions (Student's *t*-test :  $p = 0.2578$ ) or ion intensity (Student's *t*-test :  $p = 0.1806$ ) for LESA-MS/MS data (Figure 6.2), indicating that subsequent metabolite profiling with LESA-MS/MS can be performed without data loss.

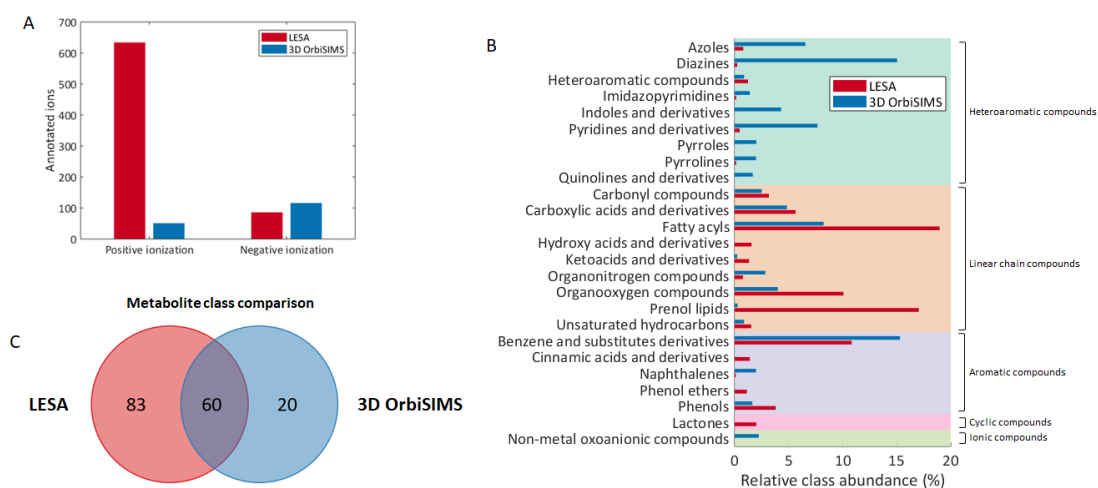


**Figure 6.2:** Investigating the effect of 3D OrbiSIMS analysis on subsequent LESA-MS/MS analysis. The signal intensity of features identified in LESA-MS spectra did not seem to be affected by prior 3D OrbiSIMS analysis. For one patient ( $n = 3$ ), the analysis was carried out with and without OrbiSIMS. The scatter plot reveals a good correlation between obtained feature intensities. Furthermore, Student's  $t$ -test revealed no significant difference in signal intensity distribution

#### 6.4.2 COMPLEMENTARY METABOLITE PROFILING WITH 3D ORBISIMS AND LESA-MS/MS

The use of direct MS has not been reported yet for ependymoma metabolite profiling due to the small quantity of tissue available per patient. Tumour sections on a TMA range in diameter from 0.6 to 1.0 mm and are approximately 4  $\mu\text{m}$  thick. 3D OrbiSIMS and LESA-MS/MS are both applicable to sub-millimetre sample size and therefore present an opportunity to obtain novel MS data from extremely small tissue quantities. Furthermore, the relatively non-destructive nature of the 3D OrbiSIMS allows for subsequent analysis of the same sample by multiple techniques. MS1 spectra acquired from the TMAs with 3D OrbiSIMS and LESA were processed to obtain ion intensities. A higher number of putative identifications (Figure 6.3A) were observed in the LESA (annotated ions: 634 (+), 51 (-)) compared to 3D OrbiSIMS data (annotated ions: 86 (+), 116 (-)). Annotated metabolites in 3D OrbiSIMS spectra were found to be mostly (hetero)aromatic compounds whilst with LESA, linear polar and non-polar

molecules were detected (Figure 6.3B). Comparison of annotated metabolite classes revealed that 60 classes were commonly detected with 3D OrbiSIMS and LESA-MS whilst 20 and 83 classes were uniquely detected by 3D OrbiSIMS and LESA-MS, respectively (Figure 6.3C). To confirm metabolite identities, LESA-MS/MS spectra were compared against a library (mzCloud). In total, 288 metabolites were identified within 4 minutes of LESA-MS/MS analysis time per sample. The number of identified metabolites is a vast improvement compared to previous research conducted on metabolite profiling of endodermoma tissue using NMR where 18 and 25 metabolites were identified and quantified, respectively. [272, 301].



**Figure 6.3:** Identifying putative metabolites in 3D OrbiSIMS and LESA-MS spectra. (A) In total, more ions were identified in the LESA-MS spectra, though the number of ions in negative mode identified as metabolite was higher for SIMS. (B) For all identified ions, the class as described in the Human Metabolome Database was obtained to identify which classes can be detected with each technique. The patched areas represent structurally similar classes. (C) A Venn diagram was derived with data from both surface MS techniques, identifying common and unique metabolite classes that could therefore provide complementary and additive information.

### 6.4.3 PREDICTING TUMOR RECURRENCE FROM METABOLITE PROFILES

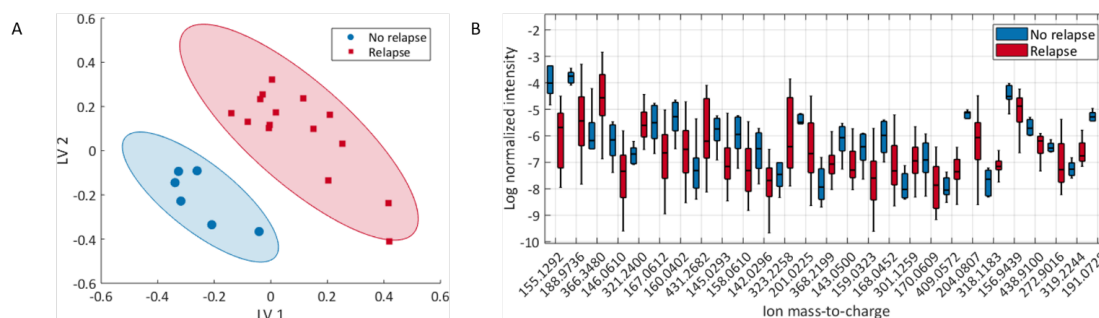
Acquired 3D OrbiSIMS and LESA-MS/MS spectra were processed using a custom MATLAB script for peak picking, alignment, data fusion and multivariate analysis. To investigate the potential of the serial MS strategy for oncometabolomics, samples

were categorised by eventual tumour recurrence. Analysed tumour sections consisted of 7 primary tumours with each three replicates, though for 5 out of 7 patients in this dataset, the tumour recurred at least once after surgery or follow-up treatment. From a diagnostic and therapeutic point of view, it is of interest to predict if patients will experience tumour relapse or if primary treatment is successful so therapy can be adjusted consequently. Waziri *et al.* [302] did not find any relation between ependymoma recurrence and demographic factors, tumor grade and location, indicating that the observed difference in metabolite profiles could be due to actual biological differences between *no relapse* and *eventual relapse* ependymoma metabolite profiles.

Fused data was subjected to PLS-DA with subsequent leave-one-out cross validation. The patient data separated into two distinct groups with recurring samples scoring more highly in both the first and second component (Figure 6.4A). The observed  $Q^2$  value (goodness-of-prediction;  $Q^2 = 0.4334$ ) for the PLS-DA model showed to be acceptable for a biological model [208].

To identify the discriminative ions between ependymoma subgroups, the VIP score for each ion was calculated. A VIP score  $\geq 1.5$  was considered discriminative. The ions which met this condition were subjected to Student's *t*-test to determine which ions were significantly altered between the two groups. In total, we identified 27 significant ions ( $p < 0.05$ ; Figure 6.4B). From those 27 ions, 18 ions could be assigned a putative molecular formula using the Human Metabolome Database [291] (Table 6.2). Most of the significant ions were found to be more prominent in the no relapse group.





**Figure 6.4:** Identification of significant ions from fused LESA-MS/MS and 3D OrbitrapSIMS data. (A) PLS-DA scores plot reveals clustering of patients based on tumour recurrence. (B) Box plot for significant ions ( $p < 0.05$ ) identified using Student's  $t$ -test and FDR estimation through a permutation test.

**Table 6.2:** Annotations for significant ions ( $p < 0.05$ ). Fold changes were calculated by dividing the average ion intensity of the *no relapse* group by the average ion intensity in the *eventual relapse* group

$m/z$	Formula	Adduct	Instrument	Fold change
321.2400	$C_{18}H_{34}O_3$	$[M+Na]^+$	LESA	0.23
145.0293	$C_9H_6O_2$	$[M+H-H_2O]^+$	LESA	1.98
201.0225	$C_8H_{10}O_5S$	$[M+H-H_2O]^+$	LESA	1.64
143.0500	$C_5H_{12}OS$	$[M+Na]^+$	LESA	1.57
170.0609	$C_{11}H_8NO_2$	$[M+H-H_2O]^+$	LESA	1.89
319.2244	$C_{18}H_{32}O_3$	$[M+Na]^+$	LESA	0.45
146.0610	$C_9H_{11}NO_2$	$[M-H-H_2O]^-$	SIMS	2.53
167.0612	$C_{11}H_8N_2$	$[M-H]^-$	SIMS	2.42
160.0402	$C_9H_9NO_3$	$[M-H-HH_2O]^-$	SIMS	2.39
145.0293	$C_9H_6O_2$	$[M-H]^-$	SIMS	1.98
158.0610	$C_{10}H_{11}NO_2$	$[M-H-H_2O]^-$	SIMS	2.25
142.0296	$C_9H_7NO_2$	$[M-H-H_2O]^-$	SIMS	1.57
201.0225	$C_8H_{10}O_4S$	$[M-H]^-$	SIMS	1.64
143.0500	$C_{10}H_8O$	$[M-H]^-$	SIMS	2.50
168.0452	$C_{11}H_9NO_2$	$[M-H-H_2O]^-$	SIMS	1.94
170.0609	$C_{11}H_{11}NO_2$	$[M-H-H_2O]^-$	SIMS	1.89

Table 6.2 (continued)

<i>m/z</i>	Formula	Adduct	Instrument	Fold change
409.0570	C <sub>21</sub> H <sub>14</sub> O <sub>9</sub>	[M-H] <sup>-</sup>	SIMS	0.44
318.1183	C <sub>13</sub> H <sub>23</sub> NO <sub>9</sub>	[M-H-H <sub>2</sub> O] <sup>-</sup>	SIMS	0.49

The putatively identified metabolites covered in total 14 genes (Table 6.3). A number of studies compared primary and recurrent tumours at transcriptomic level [302–305]. Those 14 genes identified in this study were not reported to be significantly changed between primary and recurrent ependymoma in previous transcriptomic research. This could be explained by the fact that all tumours in our study were primary from which for five patients the ependymoma eventual relapsed. From our data, we could hypothesize that the metabolite profiles of primary non-recurrent and primary recurrent ependymoma are different, but primary recurrent ependymoma do not reflect the metabolite profile of recurrent ependymoma.

**Table 6.3:** Putative metabolite IDs for significantly affected pathways ( $p < 0.05$ ) between ependymoma sub groups. Pathway analysis was performed using MetExplore

Pathway	Putative significant metabolite IDs	Related genes
Tryptophan metabolism	4,6-dihydroxyquinoline (↑)	
	5-hydroxytryptophol (↑)	ADHFE1
	β-carboline (↑)	MAOA
	Methyl indole-3-acetate (↑)	MAOB
	Quinolone-4,8,-diol (↑)	
Linoleate metabolism	12(13)-EpOME (↓)	EPHX1
	9(10)-EpOME (↓)	EPHX2
Cytochrome metabolism	Coumarin (↑)	CYP2A13
	Napthalene epoxide (↑)	CYP2A6
		CYP2F1
Phenylalanine metabolism		DDC
		GOT1
	L-phenylalanine (↑)	GOT2
		PAH
		TAT
Tyrosine metabolism	1,2-dehydrosalsolinol (↑)	
	6,7-dihydroxy-1,2,3,4-tetrahydroisoquinoline (↑)	GSTK1
	Adrenochrome (↑)	

Arrows indicate higher (↑) or lower (↓) abundance of the metabolite in the eventual relapse group compared to the no relapse group

Although only a small number of patients were used in this study, the data sup-

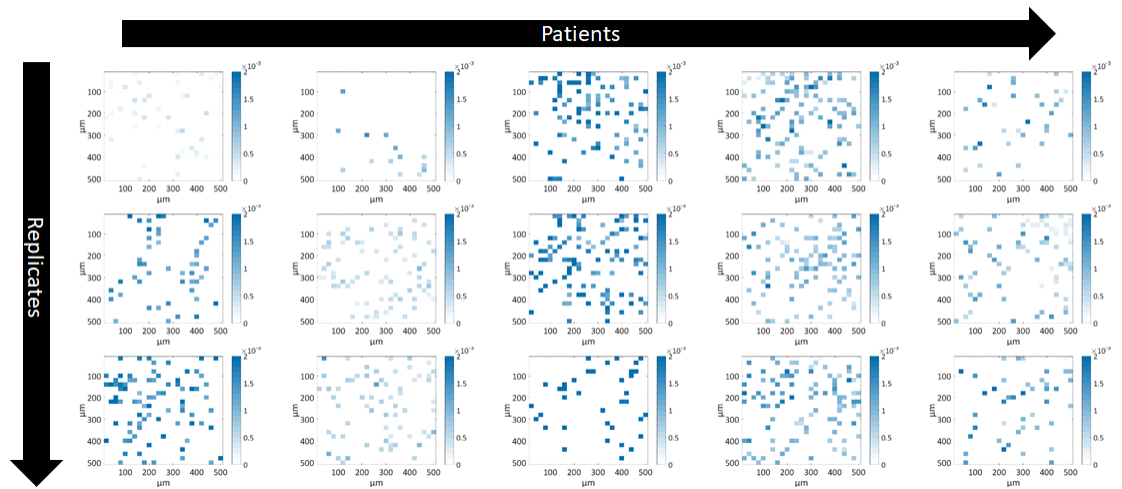
port the potential for performing large-scale studies on TMA libraries to investigate the *in situ* alterations of metabolite profiles and affected pathways between tumour phenotypes using LESA-MS and 3D OrbiSIMS. Rather than finding the same classifiers, performing analysis with both direct MS techniques allowed us to discover unique classifiers and metabolic pathways for predicting ependymoma relapse from metabolic signatures in primary ependymoma. This provides an opportunity to perform a clinical study on available libraries of FFPE tumour TMAs. Furthermore, novel pathways have been discovered that provide insight into potential tumour relapse which will improve upon understanding of tumour biology [301] and eventually leading to the discovery of new drug targets, which may prevent or delay tumour relapse.

#### **6.4.4 INVESTIGATING METABOLIC INTRA-TUMOUR HETEROGENEITY USING 3D ORBISIMS IMAGING**

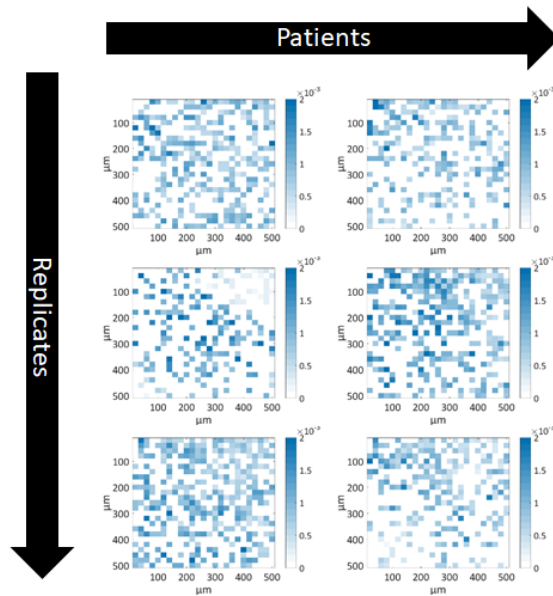
Paediatric ependymoma is characterised by a high degree of inter-patient genetic, epigenetic and metabolic heterogeneity, which manifests as distinct tumour sub-groups with varying prognoses [301, 306]. However, it is not yet known whether metabolomic profiles vary within intra-tumour regions of paediatric ependymoma and/or within distinct ecological niches.

Gularyan *et al.* [42] have shown the potential and capability of SIMS to study inter- and intra-tumour heterogeneity. Increased intra-tumour heterogeneity has been identified as a reason for treatment failure [42]. From this could be hypothesized that intra-tumour heterogeneity will increase for (eventual) recurrent tumours. Previous research has shown that SIMS can be used to investigate inter- and intra-tumour heterogeneity [42]. With the acquired 3D OrbiSIMS images, the intra-tumour heterogeneity was investigated between the *no relapse* and *eventual relapse* group. Through untargeted metabolomics, phenylalanine metabolism was putatively identified as a prominent pathway. Previous work on metabolite profiling of ependymoma showed via NMR that L-phenylalanine is highly abundant and an important discriminative metabolite for ependymoma among other paediatric brain tumours [272, 301]. The distribution of

L-phenylalanine ( $m/z$  146.0610  $[M-H-H_2O]^-$ ) across intra-tumour regions was investigated using the 3D OrbiSIMS images. Understanding the phenomena of heterogeneity is an important step towards developing personalised therapies [307]. Mass spectrometry imaging (MSI) with 3D OrbiSIMS was performed on a  $500\text{ }\mu\text{m} \times 500\text{ }\mu\text{m}$  area. In the *eventual relapse* group (Figure 6.5), the presence of L-phenylalanine appears to be lower across the analysed area compared to the *no relapse* group (Figure 6.6).

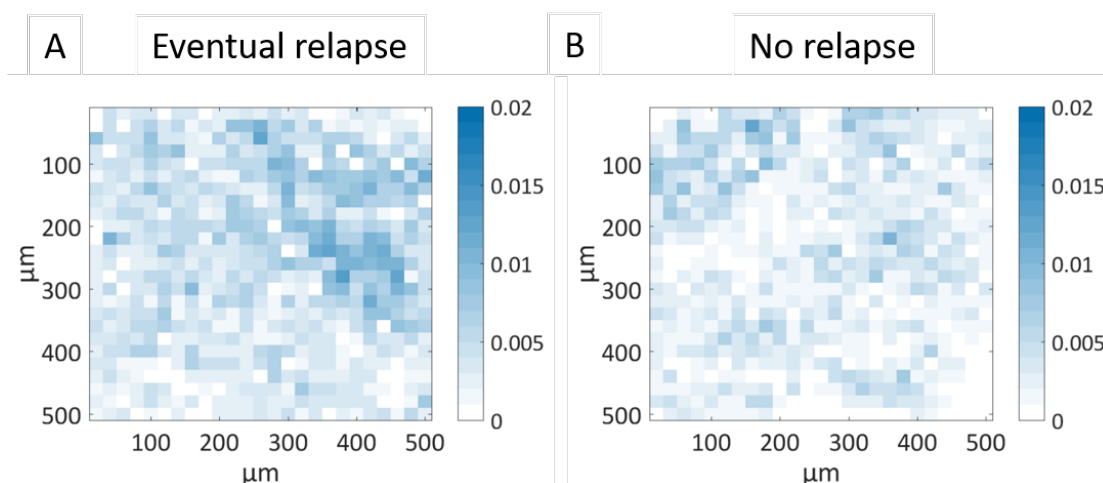


**Figure 6.5:** Ion images for L-phenylalanine ( $m/z$  146.0610  $\pm$  0.001  $m/z$ ) for patients in the *eventual relapse* group

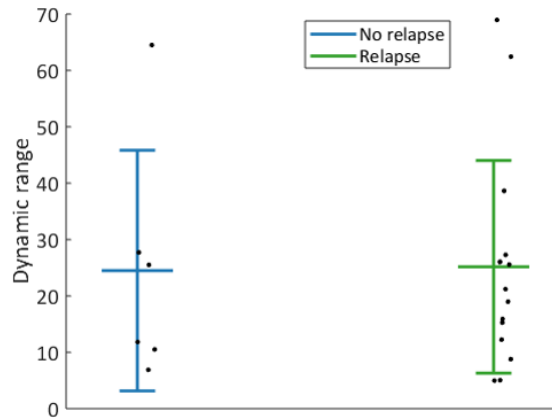


**Figure 6.6:** Ion images for L-phenylalanine ( $m/z$  146.0610  $\pm$  0.001  $m/z$ ) for patients in the *no relapse* group

To investigate the intra-tumour heterogeneity of L-phenylalanine between the *no relapse* and *eventual relapse* group, the dynamic range (ratio between maximum ion and minimum ion intensity) was used as a measure [295]. For comparison of the degree of heterogeneity, adenine ( $m/z$  134.0467 [M-H]<sup>−</sup>) was used because of its spatial abundance (presence across the analysed area; Figure 6.7A-B). The dynamic range for adenine was found to be smaller (more homogeneous) compared to L-phenylalanine for both tumour subgroups (Levene's test:  $p = 0.0039$ ). Increased tumour heterogeneity has previously been identified as an important factor for therapy resistance and poor overall survival [308–310]. However, for L-phenylalanine, no difference in intra-tumour heterogeneity was observed between both tumour groups (Figure 6.8). This could be a potential indicator that intra-tumour heterogeneity in primary ependymoma does not have a predictive value for eventual relapse. In addition, this could be an indication that increased chemical heterogeneity occurs as a result of relapse and cannot be observed in primary tumour subgroups. However, this would require investigation of a larger sample cohort to confirm current findings.

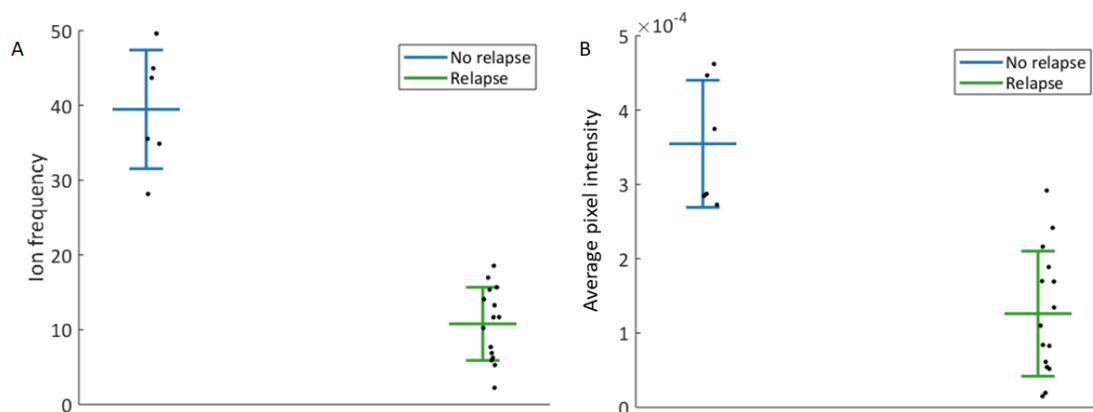


**Figure 6.7:** Comparison of intra-tumour heterogeneity between the *no relapse* and *eventual relapse* subgroups. Representative 3D OrbiSIMS images for adenine in the (A) *eventual relapse* and (B) *no relapse* group. The dynamic range for adenine was significantly smaller compared to L-phenylalanine (ANOVA:  $p = 0.0039$ ).



**Figure 6.8:** Investigating of difference in intra-tumour heterogeneity between *no relapse* and *eventual relapse* subgroups. No difference in heterogeneity (dynamic range) was observed.

Despite of no difference in intra-tumour heterogeneity, L-phenylalanine seems to be less present (higher number of zero intensity pixels) and less abundant in the *eventual relapse* subgroup (Figure 6.5). Statistical comparison revealed that the spatial presence of L-phenylalanine is significantly higher in the *no relapse* group (Mann-Whitney:  $p = 5.29 \times 10^{-4}$ ) Furthermore, it was found that, on average, the abundance (non-zero pixel intensity) for L-phenylalanine was higher in the *no relapse* group. Based on these results, the spatial absence of L-phenylalanine seems to be a signature for eventual ependymoma relapse. Nonetheless, a larger sample cohort is required to validate these findings .



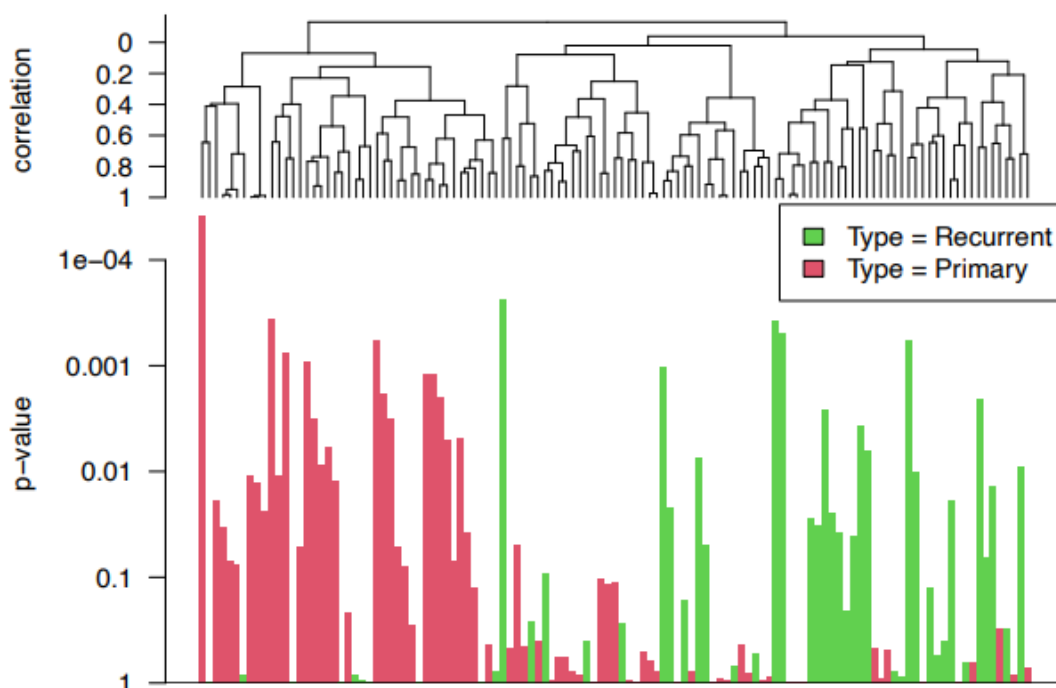
**Figure 6.9:** Statistical comparison for the spatial abundance of L-phenylalanine between ependymoma subgroups. (A) The number of pixels (ion frequency) in which L-phenylalanine was detected was higher for the *no relapse* group (Mann-Whitney:  $p = 5.29 \times 10^{-4}$ ). (B) In addition, the average pixel intensity for L-phenylalanine was found to be higher in the *no relapse* group (Student's *t*-test :  $p = 2.12 \times 10^{-5}$ )

#### 6.4.5 VALIDATION OF MS-BASED METABOLOMICS WITH PUBLICLY AVAILABLE GENE EXPRESSION DATA

Due to the limited availability of ependymoma samples, only a small cohort could be analyzed. Through pathway analysis, it was found that fourteen genes are affected (differential expression) between *no relapse* and *eventual relapse* ependymoma cohorts. To validate these findings, publicly available gene expression datasets were used for primary ( $n = 72$ ) and recurrent ( $n = 47$ ) paediatric ependymoma [296, 297]. Ten out of the fourteen genes listed in Table 3 were present in these gene expression datasets. From those 10 genes, four genes showed a significant differential expression between the ependymoma subgroups. ADHFE1 ( $p = 0.00234$ ) was upregulated in primary ependymoma, whilst GSTK1 ( $p = 0.00450$ ), GOT2 ( $p = 0.0393$ ) and EPHX2 ( $p = 0.0433$ ) showed a significantly higher expression in the recurrent ependymoma cohort. The gene expression results are partially in concordance with the metabolomics data. The expression of GSTK1 and GOT2 is in line with higher abundance of 1,2-dehydrosalsolinol-6,7-dihydroxy-1,2,3,4-tetrahydroisoquinoline, adrenochrome, and L-phenylalanine, respectively. On the other hand, the expression of ADHFE1 and



EPHX2 was opposite to the observed difference in metabolite abundance between the *no relapse* and *eventual relapse* subgroups. An explanation could be that for the metabolomics study only primary ependymomas were used and these might not completely reflect the same gene expression profile as recurrent ependymoma. This would require further investigation using a larger sample cohort. In addition, the group of ten genes as a single gene set showed a significant change between primary and recurrent ependymoma (global test:  $p = 0.00257$ ) and showed clear separation of primary tumours from recurrent ependymal tumours (Figure 6.10), indicating a good predictive power of the four significant genes for predicting ependymoma relapse. Although our metabolomics dataset is small, the gene expression analysis supports the significance of the genes identified through MS-based metabolomics and pathway analysis. Moreover, excellent clustering was observed for primary and recurrent ependymoma based on the subset of the significant genes. This confirmation of the metabolomics data shows the potential of the sequential MS strategy to be further used for large scale clinical studies on archived TMAs.



**Figure 6.10:** Gene expression analysis of primary ( $n = 72$ ) and recurrent ependymoma ( $n = 47$ ). The dendrogram reveals two distinct groups for which most of the recurrent samples belong to the first group and most of the primary samples to the second group. The samples associated with strong evidence for the association between the response (primary vs. recurrent) and the gene expression profile of the gene set (10 genes) have small p-values (tall bars in the bottom plot)

## 6.5 CONCLUSION

In this chapter, a novel mass spectrometry strategy is presented for metabolite profiling of tumour microarrays. Complementary metabolite profiles were obtained permitting putative identification of additional affected metabolic pathways and their corresponding genes, resulting in a putative predictive signature of no relapse/relapse. This opens new opportunities to perform large scale metabolomics studies on archived tissue libraries. Furthermore, the minimally required sample preparation and short analysis time (10 minutes with 3D OrbiSIMS; 4 minutes with LESA-MS/MS) permits high sample throughput, making this strategy a competitive alternative to standard

## CHAPTER 6

metabolomics analysis such as GC-MS, LC-MS and NMR.

## Overall conclusion & future perspectives

The aim of this thesis was to develop a LESA-MS/MS protocol for high throughput quantitative analysis of proteins adsorbed onto synthetic polymer surfaces. Robust surface extraction was achieved through the use of superhydrophobic-superhydrophilic patterned arrays (Droplet Microarray) which were also found to be suitable as printing substrate for monomer solutions and fabrication of LESA-compatible polymer microarrays. LESA-MS/MS parameters were optimised through a series of experiments in order to obtain a stable spray signal and robust identification of peptides. Furthermore, *in situ* digestion conditions were optimised to achieve reproducible results. The developed method was then tested on a selection of polymers in array as well as scaled-up (6-well plate) format. The quantitative results showed a nonlinear trend between cell response (attachment) and protein adsorption, i.e. at a certain level of protein adsorption the increase cell response is driven by another unknown factor. This shows the potential of the developed analysis strategy to be further used for screening polymer libraries in order to increase the understanding of cell-material and protein-material interactions.

In the current setup, a fully-defined and cell-free medium (Essential 8™) was used to outline the analysis strategy. The next step would be to increase the complexity of the culture medium by adding cells and investigate what the composition of the

secreted matrix is. This will provide with additional information for understanding the cell response mechanism on synthetic polymer surfaces. However, this would require the development a protocol the remove cells from the surface without causing cell lysis and/or removal of the deposited matrix. Currently, no protocol has been available to achieve this. In addition, it should also be explored if the method can handle more complex matrices. This could be done through repeating the array screening experiments with a more complex cell culture medium, e.g. a culture medium supplemented with serum proteins. This would require first full characterisation of the culture medium through LC-MS/MS to create a library of the composition before starting LESA-MS/MS analysis. The results of the LESA-MS/MS could then be compared to the LC-MS/MS results to investigate what information is lost through LESA-MS/MS analysis. Furthermore, increased complexity of the proteome might require revisiting the used quantification strategy. One could think if introducing (isotopically-labelled) internal standards, though, experiments can get for costly and so far no study has used *in situ* labelling for quantification of proteins. Therefore, a set of experiments would be required to investigate how to introduce an internal standard for proper normalisation of technical variation.

Besides the use of this quantitative LESA-MS/MS method in the field of biomaterials discovery, it might also be of interest for use in the clinical field. For instance, the study performed in Chapter 6 could be repeated using the developed protocol the find differences in the proteome of ependymoma subgroups. This could provide with further understanding of tumour progression, relapse, survival chances, and identification of targets for therapy.

# Bibliography

- [1] William M Gallagher, Iseult Lynch, Lorcan T Allen, Ian Miller, Stephen C Penney, Darran P O'Connor, Stephen Pennington, Alan K Keenan, and Kenneth A Dawson. Molecular basis of cell–biomaterial interaction: Insights gained from transcriptomic and proteomic studies. *Biomaterials*, 27(35):5871–5882, 2006.
- [2] Andrew L Hook, Chien-Yi Chang, Jing Yang, Jeni Luckett, Alan Cockayne, Steve Atkinson, Ying Mei, Roger Bayston, Derek J Irvine, Robert Langer, Daniel G Anderson, Paul Williams, Martyn C Davies, and Morgan R Alexander. Combinatorial discovery of polymers resistant to bacterial attachment. *Nature Biotechnology*, 30(9):868–875, sep 2012. ISSN 1087-0156. doi: 10.1038/nbt.2316.
- [3] Davide Campoccia, Lucio Montanaro, and Carla Renata Arciola. A review of the biomaterials technologies for infection-resistant surfaces. *Biomaterials*, 34(34): 8533–8554, 2013.
- [4] WSW Harun, RIM Asri, J Alias, FH Zulkifli, K Kadirgama, SAC Ghani, and JHM Shariffuddin. A comprehensive review of hydroxyapatite-based coatings adhesion on metallic biomaterials. *Ceramics International*, 44(2):1250–1268, 2018.
- [5] Jason A. Burdick, Laney M. Philpott, and Kristi S. Anseth. Synthesis and characterization of tetrafunctional lactic acid oligomers: A potential in situ forming degradable orthopaedic biomaterial. *Journal of Polymer Science Part A: Polymer Chemistry*, 39(5):683–692, mar 2001. ISSN 1099-0518. doi: 10.1002/1099-0518(20010301)39:5<683::AID-POLA1040>3.0.CO;2-Z.

- [6] Anthony D. Metcalfe and Mark W.J. Ferguson. Tissue engineering of replacement skin: The crossroads of biomaterials, wound healing, embryonic development, stem cells and regeneration, jun 2007. ISSN 17425662.
- [7] Kalpana S. Katti. Biomaterials in total joint replacement. *Colloids and Surfaces B: Biointerfaces*, 39(3):133–142, dec 2004. ISSN 09277765. doi: 10.1016/j.colsurfb.2003.12.002.
- [8] Darya Hadavi and André A. Poot. Biomaterials for the treatment of Alzheimer’s disease, jun 2016. ISSN 22964185.
- [9] Jaesung Kim, Pyung Hwan Kim, Sung Wan Kim, and Chae Ok Yun. Enhancing the therapeutic efficacy of adenovirus in combination with biomaterials, feb 2012. ISSN 01429612.
- [10] Zachary S. Dunn, John Mac, and Pin Wang. T-cell immunotherapy enhanced by designer biomaterials. *Biomaterials*, 217:119265, 2019. ISSN 18785905. doi: 10.1016/j.biomaterials.2019.119265.
- [11] Adam A. Dundas, Olutoba Sanni, Jean-Frédéric Dubern, Georgios Dimitrakis, Andrew L. Hook, Derek J. Irvine, Paul Williams, and Morgan R. Alexander. Validating a Predictive Structure–Property Relationship by Discovery of Novel Polymers which Reduce Bacterial Biofilm Formation. *Advanced Materials*, 31(49): 1903513, dec 2019. ISSN 0935-9648. doi: 10.1002/adma.201903513.
- [12] Marjo Yliperttula, Bong Geun Chung, Akshay Navaladi, Amir Manbachi, and Arto Urtti. High-throughput screening of cell responses to biomaterials. *European Journal of Pharmaceutical Sciences*, 35(3):151–160, oct 2008. ISSN 09280987. doi: 10.1016/j.ejps.2008.04.012.
- [13] Carl G. Simon and Lin Gibson Sheng. Combinatorial and high-throughput screening of biomaterials. *Advanced Materials*, 23(3):369–387, jan 2011. ISSN 09359648. doi: 10.1002/adma.201001763.

- [14] Joachim Kohn. New approaches to biomaterials design. *Nature Materials*, 3(11): 745–747, 2004. ISSN 14761122. doi: 10.1038/nmat1249.
- [15] Mohammed S Algahtani, David J Scurr, Andrew L Hook, Daniel G Anderson, Robert S Langer, Jonathan C BNOOPurley, Morgan R Alexander, and Martyn C Davies. High throughput screening for biomaterials discovery. *Journal of Controlled Release*, 190:115–126, 2014.
- [16] Andrew L. Hook, Chien-Yi Chang, Jing Yang, Steve Atkinson, Robert Langer, Daniel G. Anderson, Martyn C. Davies, Paul Williams, and Morgan R. Alexander. Discovery of Novel Materials with Broad Resistance to Bacterial Attachment Using Combinatorial Polymer Microarrays. *Advanced Materials*, 25(18):2542–2547, may 2013. ISSN 09359648. doi: 10.1002/adma.201204936.
- [17] Jeffrey A. Hubbell. Biomaterials science and high-throughput screening. *Nature Biotechnology*, 22(7):828–829, jul 2004. ISSN 10870156. doi: 10.1038/nbt0704-828.
- [18] Asha K. Patel, Mark W. Tibbitt, Adam D. Celiz, Martyn C. Davies, Robert Langer, Chris Denning, Morgan R. Alexander, and Daniel G. Anderson. High throughput screening for discovery of materials that control stem cell fate. *Current Opinion in Solid State and Materials Science*, 20(4):202–211, aug 2016. ISSN 13590286. doi: 10.1016/j.cossms.2016.02.002.
- [19] Andrew L. Hook, Philip M. Williams, Morgan R. Alexander, and David J. Scurr. Multivariate ToF-SIMS image analysis of polymer microarrays and protein adsorption. *Biointerphases*, 10(1):019005, mar 2015. ISSN 1934-8630. doi: 10.1116/1.4906484.
- [20] Andrew L. Hook, Helmut Thissen, and Nicolas H. Voelcker. Advanced Substrate Fabrication for Cell Microarrays. *Biomacromolecules*, 10(3):573–579, mar 2009. ISSN 1525-7797. doi: 10.1021/bm801217n.
- [21] Guilhem Tourniaire, Jane Collins, Sara Campbell, Hitoshi Mizomoto, Shuichiro Ogawa, Jean-Francois Thaburet, Mark Bradley, I. Blaga, D. L. Barker, and S. B.



- Jovanovich. Polymer microarrays for cellular adhesion. *Chemical Communications*, 10(20):2118, 2006. ISSN 1359-7345. doi: 10.1039/b602009g.
- [22] Daehwan Jang, Dongjo Kim, and Jooho Moon. Influence of fluid physical properties on ink-jet printability. *Langmuir*, 25(5):2629–2635, mar 2009. ISSN 07437463. doi: 10.1021/la900059m.
- [23] Qiang Zheng, Jiangang Lu, Hao Chen, Lei Huang, Jin Cai, and Zhinan Xu. Application of inkjet printing technique for biological material delivery and antimicrobial assays. *Analytical Biochemistry*, 410(2):171–176, mar 2011. ISSN 0003-2697. doi: 10.1016/J.AB.2010.10.024.
- [24] Andrew L. Hook, Daniel G. Anderson, Robert Langer, Paul Williams, Martyn C. Davies, and Morgan R. Alexander. High throughput methods applied in biomaterial development and discovery. *Biomaterials*, 31(2):187–198, jan 2010. ISSN 01429612. doi: 10.1016/j.biomaterials.2009.09.037.
- [25] Emine Tekin, Berend-Jan de Gans, Ulrich S. Schubert, G. Huber, S. R. Nagel, and T. A. Witten. Ink-jet printing of polymers - From single dots to thin film libraries. *Journal of Materials Chemistry*, 14(17):2627, aug 2004. ISSN 0959-9428. doi: 10.1039/b407478e.
- [26] Tiago G. Fernandes, Maria Margarida Diogo, Douglas S. Clark, Jonathan S. Dordick, and Joaquim M.S. S Cabral. High-throughput cellular microarray platforms: applications in drug discovery, toxicology and stem cell research. *Trends in Biotechnology*, 27(6):342–349, jun 2009. ISSN 01677799. doi: 10.1016/j.tibtech.2009.02.009.
- [27] Daniel G. Anderson, David Putnam, Erin B. Lavik, Tahir A. Mahmood, and Robert Langer. Biomaterial microarrays: rapid, microscale screening of polymer–cell interaction. *Biomaterials*, 26(23):4892–4897, aug 2005. ISSN 0142-9612. doi: 10.1016/J.BIOMATERIALS.2004.11.052.

- [28] Jeffrey A. Hubbell. Biomaterials in Tissue Engineering. *Nature Biotechnology*, 13 (6):565–576, jun 1995. ISSN 1087-0156. doi: 10.1038/nbt0695-565.
- [29] Christopher J Flaim, Shu Chien, and Sangeeta N Bhatia. An extracellular matrix microarray for probing cellular differentiation. *Nature Methods*, 2(2):119–125, feb 2005. ISSN 1548-7091. doi: 10.1038/nmeth736.
- [30] Yoav Soen, Akiko Mori, Theo D Palmer, and Patrick O Brown. Exploring the regulation of human neural precursor cell differentiation using arrays of signaling microenvironments. *Molecular Systems Biology*, 2(1):37, jan 2006. ISSN 1744-4292. doi: 10.1038/msb4100076.
- [31] Masafumi Nakajima, Toshinari Ishimuro, Koichi Kato, In-Kap Ko, Isao Hirata, Yusuke Arima, and Hiroo Iwata. Combinatorial protein display for the cell-based screening of biomaterials that direct neural stem cell differentiation. *Biomaterials*, 28(6):1048–1060, feb 2007. ISSN 01429612. doi: 10.1016/j.biomaterials.2006.10.004.
- [32] Ying Mei, Krishanu Saha, Said R. Bogatyrev, Jing Yang, Andrew L. Hook, Z. Ilke Kalcioğlu, Seung-Woo Cho, Maisam Mitalipova, Neena Pyzocha, Fredrick Rojas, Krystyn J. Van Vliet, Martyn C. Davies, Morgan R. Alexander, Robert Langer, Rudolf Jaenisch, and Daniel G. Anderson. Combinatorial development of biomaterials for clonal growth of human pluripotent stem cells. *Nature Materials*, 9 (9):768–778, sep 2010. ISSN 1476-1122. doi: 10.1038/nmat2812.
- [33] Adam D. Celiz, James G. W. Smith, Asha K. Patel, Andrew L. Hook, Divya Rajamohan, Vinoj T. George, Luke Flatt, Minal J. Patel, Vidana C. Epa, Taranjit Singh, Robert Langer, Daniel G. Anderson, Nicholas D. Allen, David C. Hay, David A. Winkler, David A. Barrett, Martyn C. Davies, Lorraine E. Young, Chris Denning, and Morgan R. Alexander. Discovery of a Novel Polymer for Human Pluripotent Stem Cell Expansion and Multilineage Differentiation. *Advanced Materials*, 27(27):4006–4012, jul 2015. ISSN 09359648. doi: 10.1002/adma.201501351.

- [34] J. C. Vickerman and D. Briggs. *ToF-SIMS: materials analysis by mass spectrometry*. IMPublications Open, IM Publica edition, 2013.
- [35] Sadia Rabbani, Andrew M. Barber, John S. Fletcher, Nicholas P. Lockyer, and John C. Vickerman. TOF-SIMS with argon gas cluster ion beams: A comparison with C 60+. *Analytical Chemistry*, 83(10):3793–3800, may 2011. ISSN 00032700. doi: 10.1021/ac200288v.
- [36] Greg Gillen and Albert Fahey. Secondary ion mass spectrometry using cluster primary ion beams. In *Applied Surface Science*, volume 203-204, pages 209–213. Elsevier, jan 2003. doi: 10.1016/S0169-4332(02)00627-X.
- [37] Nicholas J. Popczun, Lars Breuer, Andreas Wucher, and Nicholas Winograd. On the SIMS Ionization Probability of Organic Molecules. *Journal of The American Society for Mass Spectrometry*, 28(6):1182–1191, jun 2017. ISSN 1044-0305. doi: 10.1007/s13361-017-1624-0.
- [38] N. Kh Dzhemilev, U. Kh Rasulev, and S. V. Verkhoturov. The fragmentation of sputtered cluster ions and their contribution to secondary ion mass-spectra. *Nuclear Inst. and Methods in Physics Research, B*, 29(3):531–536, dec 1987. ISSN 0168583X. doi: 10.1016/0168-583X(87)90065-6.
- [39] A.J Urquhart, M Taylor, D.G Anderson, R Langer, M.C Davies, and M.R Alexander. TOF-SIMS Analysis of a 576 Micropatterned Copolymer Array To Reveal Surface Moieties That Control Wettability. *Analytical Chemistry*, 80(1):135–142, 2007. doi: 10.1021/AC071560K.
- [40] Helmut Liebl. Secondary- ion mass spectrometry and its use in depth profiling. *Journal of Vacuum Science and Technology*, 12(1):385–391, 1975.
- [41] Melissa K Passarelli, Alexander Pirkel, Rudolf Moellers, Dmitry Grinfeld, Felix Kollmer, Rasmus Havelund, Carla F Newman, Peter S Marshall, Henrik Arlinghaus, Morgan R Alexander, Andy West, Stevan Horning, Ewald Niehuis, Alexander Makarov, Colin T Dollery, and Ian S Gilmore. The 3D OrbiSIMS—label-free

- metabolic imaging with subcellular lateral resolution and high mass-resolving power. *Nature Methods*, 14(12):1175–1183, nov 2017. ISSN 1548-7091. doi: 10.1038/nmeth.4504.
- [42] Samvel K Gularyan, Alexander A Gulin, Ksenia S Anufrieva, Victoria Shender, Michail I Shakhparonov, Soniya Bastola, Nadezhda V. Antipova, Tatiana F Kovalenko, Yury P Rubtsov, Yaroslav A Latyshev, Alexander A Potapov, and Marat S Pavlyukov. Investigation of inter- and intra-tumoral heterogeneity of glioblastoma using TOF-SIMS. *Molecular & Cellular Proteomics*, page mcp.RA120.001986, apr 2020. ISSN 1535-9476. doi: 10.1074/mcp.RA120.001986.
- [43] Melissa K. Passarelli and Nicholas Winograd. Lipid imaging with time-of-flight secondary ion mass spectrometry (ToF-SIMS). *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, 1811(11):976–990, nov 2011. ISSN 13881981. doi: 10.1016/j.bbalip.2011.05.007.
- [44] J.R. Ferraro and K. Nakamoto. *Introductory Raman Spectroscopy*. Elsevier, 2003.
- [45] Douglas A Skoog, F James Holler, and Stanley R Crouch. *Principles of instrumental analysis*. Cengage learning, 2017.
- [46] D. R. Baer and S. Thevuthasan. Characterization of Thin Films and Coatings. In *Handbook of Deposition Technologies for Films and Coatings*, pages 749–864. Elsevier Inc., jan 2010. ISBN 9780815520313. doi: 10.1016/B978-0-8155-2031-3.00016-8.
- [47] P. Van Der Heide. *X-ray Photoelectron Spectroscopy: An Introduction to Principles and Practices*. John Wiley & Sons, Inc., 2012.
- [48] CJ Powell and A Jablonski. Surface sensitivity of x-ray photoelectron spectroscopy. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 601(1-2):54–65, 2009.
- [49] M. H. Engelhard, T. C. Droubay, and Y. Du. X-ray photoelectron spectroscopy

- applications. In *Encyclopedia of Spectroscopy and Spectrometry*, pages 716–724. Elsevier, jan 2016. ISBN 9780128032244. doi: 10.1016/B978-0-12-409547-2.12102-X.
- [50] Tommi Huhtamäki, Xuelin Tian, Juuso T. Korhonen, and Robin H.A. Ras. Surface-wetting characterization using contact-angle measurements. *Nature Protocols*, 13(7):1521–1538, jul 2018. ISSN 17502799. doi: 10.1038/s41596-018-0003-z.
- [51] Colin D Bain and George M Whitesides. A study by contact angle of the acid-base behavior of monolayers containing. omega.-mercaptocarboxylic acids adsorbed on gold: an example of reactive spreading. *Langmuir*, 5(6):1370–1378, 1989.
- [52] By Thomas Young and M D For Sec. III. An essay on the cohesion of fluids. *Philosophical Transactions of the Royal Society of London*, 95:65–87, dec 1805. ISSN 0261-0523. doi: 10.1098/rstl.1805.0005.
- [53] I.T. Jolliffe. Principal Component Analysis and Factor Analysis. In *Principal Component Analysis*, pages 150–166. Springer-Verlag, New York, 2002. doi: 10.1007/0-387-22440-8\_7.
- [54] D L Massart, B G M Vandeginste, L M C Buydens, S De Jong, P J Lewi, and J Smeyers-Verbeke. Principal components. In *Handbook of Chemometrics & Qualimetrics: Part A*, chapter 17, pages 519–556. Elsevier, 1 edition, 1998. doi: 10.1016/S0922-3487(97)80047-0.
- [55] Hervé Abdi and Lynne J. Williams. Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(4):433–459, jul 2010. ISSN 19395108. doi: 10.1002/wics.101.
- [56] Turgay Celik. Unsupervised Change Detection in Satellite Images Using Principal Component Analysis and k-Means Clustering. *IEEE Geoscience and Remote Sensing Letters*, 6(4):772–776, oct 2009. ISSN 1545-598X. doi: 10.1109/LGRS.2009.2025059.

- [57] M. S. Wagner and David G. Castner. Characterization of Adsorbed Protein Films by Time-of-Flight Secondary Ion Mass Spectrometry with Principal Component Analysis. *Langmuir*, 17(15):4649–4660, 2001. doi: 10.1021/LA001209T.
- [58] Max Bylesjö, Mattias Rantalainen, Olivier Cloarec, Jeremy K. Nicholson, Elaine Holmes, and Johan Trygg. OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *Journal of Chemometrics*, 20(8-10):341–351, aug 2006. ISSN 08869383. doi: 10.1002/cem.1006.
- [59] Aishah Nasir, Jordan Thorpe, Laurence Burroughs, Joris Meurs, Sara Pijuan-Galito, Derek J Irvine, Morgan R Alexander, and Chris Denning. Discovery of a novel polymer for xeno-free, long-term culture of human pluripotent stem cell expansion. *Advanced Healthcare Materials*, page 2001448, 2021. doi: 10.1002/adhm.202001448.
- [60] Michael D Abràmoff, Paulo J Magalhães, and Sunanda J Ram. Image processing with imagej. *Biophotonics International*, 11(7):36–42, 2004.
- [61] Michael R. Lamprecht, David M. Sabatini, and Anne E. Carpenter. CellProfiler™: Free, versatile software for automated biological image analysis. *BioTechniques*, 42(1):71–75, jan 2007. ISSN 07366205. doi: 10.2144/000112257.
- [62] Jan Kapuscinski. Dapi: a dna-specific fluorescent probe. *Biotechnic & Histochemistry*, 70(5):220–233, 1995.
- [63] Yuin-Han Loh, Qiang Wu, Joon-Lin Chew, Vinsensius B Vega, Weiwei Zhang, Xi Chen, Guillaume Bourque, Joshy George, Bernard Leong, Jun Liu, et al. The oct4 and nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nature Genetics*, 38(4):431–440, 2006.
- [64] Timothy D Jones, Thomas M Ulbright, John N Eble, Lee Ann Baldridge, and Liang Cheng. Oct4 staining in testicular tumors: a sensitive and specific marker for seminoma and embryonal carcinoma. *The American journal of surgical pathology*, 28(7):935–940, 2004.

- [65] Ulrich Hersel, Claudia Dahmen, and Horst Kessler. RGD modified polymers: biomaterials for stimulated cell adhesion and beyond. *Biomaterials*, 24(24):4385–4415, 2003.
- [66] Ying Mei, Sharon Gerecht, Michael Taylor, Andrew J. Urquhart, Said R. Bogatyrev, Seung-Woo Cho, Martyn C. Davies, Morgan R. Alexander, Robert S. Langer, and Daniel G. Anderson. Mapping the Interactions among Biomaterials, Adsorbed Proteins, and Human Embryonic Stem Cells. *Advanced Materials*, 21(27):2781–2786, jul 2009. ISSN 09359648. doi: 10.1002/adma.200803184.
- [67] Moamen Hammad, Wei Rao, James G. W. Smith, Daniel G. Anderson, Robert Langer, Lorraine E. Young, David A. Barrett, Martyn C. Davies, Chris Denning, and Morgan R. Alexander. Identification of polymer surface adsorbed proteins implicated in pluripotent human embryonic stem cell expansion. *Biomater. Sci.*, 4(9):1381–1391, 2016. ISSN 2047-4830. doi: 10.1039/C6BM00214E.
- [68] Natália M. Alves, Iva Pashkuleva, Rui L. Reis, and João F. Mano. Controlling cell behavior through the design of polymer surfaces. *Small*, 6(20):2208–2220, oct 2010. ISSN 16136810. doi: 10.1002/sml.201000233.
- [69] Mohamed-Nur Abdallah, Simon D. Tran, Ghada Abughanam, Marco Laurenti, David Zuanazzi, Mohamed A. Mezour, Yizhi Xiao, Marta Cerruti, Walter L. Siqueira, and Faleh Tamimi. Biomaterial surface proteomic signature determines interaction with epithelial cells. *Acta Biomaterialia*, 54:150–163, mar 2017. ISSN 17427061. doi: 10.1016/j.actbio.2017.02.044.
- [70] Cameron J. Wilson, Richard E. Clegg, David I. Leavesley, and Mark J. Pearcy. Mediation of biomaterial-cell interactions by adsorbed proteins: A review. *Tissue Engineering*, 11(1-2):1–18, jan 2005. ISSN 10763279. doi: 10.1089/ten.2005.11.1.
- [71] Hynda K. Kleinman and George R. Martin. Matrigel: Basement membrane matrix with biological activity. *Seminars in Cancer Biology*, 15(5 SPEC. ISS.):378–386, oct 2005. ISSN 1044579X. doi: 10.1016/j.semcan.2005.05.004.

- [72] Nick Huettner, Tim R. Dargaville, and Aurelien Forget. Discovering Cell-Adhesion Peptides in Tissue Engineering: Beyond RGD, apr 2018. ISSN 18793096.
- [73] Nathan D. Gallant, Kristopher A. Lavery, Eric J. Amis, and Matthew L. Becker. Universal gradient substrates for "click" biofunctionalization. *Advanced Materials*, 19(7):965–969, apr 2007. ISSN 09359648. doi: 10.1002/adma.200602221.
- [74] Matt J. Kipper, Hynda K. Kleinman, and Francis W. Wang. Covalent surface chemistry gradients for presenting bioactive peptides. *Analytical Biochemistry*, 363(2):175–184, apr 2007. ISSN 00032697. doi: 10.1016/j.ab.2007.01.036.
- [75] Matt J. Kipper, Hynda K. Kleinman, and Francis W. Wang. New method for modeling connective-tissue cell migration: Improved accuracy on motility parameters. *Biophysical Journal*, 93(5):1797–1808, 2007. ISSN 00063495. doi: 10.1529/biophysj.106.096800.
- [76] Ruedi Aebersold and Matthias Mann. Mass spectrometry-based proteomics. *Nature*, 422(6928):198–207, mar 2003. ISSN 0028-0836. doi: 10.1038/nature01511.
- [77] W Y Tong, Y M Liang, V Tam, H K Yip, Y T Kao, K M C Cheung, K W K Yeung, and Y W Lam. Biochemical Characterization of the Cell-Biomaterial Interface by Quantitative Proteomics. *Molecular & Cellular Proteomics*, 9(10):2089–2098, 2010. doi: 10.1074/mcp.M110.001966.
- [78] Paul J Taylor. Matrix effects: the achilles heel of quantitative high-performance liquid chromatography–electrospray–tandem mass spectrometry. *Clinical biochemistry*, 38(4):328–334, 2005.
- [79] Ileana R. Leon, Veit Schwammle, Ole N. Jensen, and Richard R. Sprenger. Quantitative assessment of in-solution digestion efficiency identifies optimal protocols for unbiased protein analysis. *Molecular and Cellular Proteomics*, 12(10):2992–3005, oct 2013. ISSN 15359476. doi: 10.1074/mcp.M112.025585.



- [80] Zachery R. Gregorich, Ying Hua Chang, and Ying Ge. Proteomics in heart failure: Top-down or bottom-up? *Pflugers Archiv European Journal of Physiology*, 466(6): 1199–1209, mar 2014. ISSN 14322013. doi: 10.1007/s00424-014-1471-9.
- [81] Adam D. Catherman, Owen S. Skinner, and Neil L. Kelleher. Top Down proteomics: Facts and perspectives, mar 2014. ISSN 10902104.
- [82] Masaaki Oyama, Chiharu Itagaki, Hiroko Hata, Yutaka Suzuki, Tomonori Izumi, Tohru Natsume, Toshiaki Isobe, and Sumio Sugano. Analysis of small human proteins reveals the translation of upstream open reading frames of mrnas. *Genome research*, 14(10b):2048–2052, 2004.
- [83] Zhixia Zhong, Katherine L Wilson, and Kris Noel Dahl. Beyond lamins: other structural components of the nucleoskeleton. In *Methods in cell biology*, volume 98, pages 97–119. Elsevier, 2010.
- [84] Joscelyn Sarsby, Nicholas J. Martin, Patricia F. Lalor, Josephine Bunch, and Helen J. Cooper. Top-Down and Bottom-Up Identification of Proteins by Liquid Extraction Surface Analysis Mass Spectrometry of Healthy and Diseased Human Liver Tissue. *Journal of The American Society for Mass Spectrometry*, 25(11):1953–1961, nov 2014. ISSN 1044-0305. doi: 10.1007/s13361-014-0967-z.
- [85] Lijuan Kang, Naidong Weng, and Wenying Jian. LC–MS bioanalysis of intact proteins and peptides. *Biomedical Chromatography*, 34(1), jan 2020. ISSN 0269-3879. doi: 10.1002/bmc.4633.
- [86] Wei Sun, Shuzhen Wu, Xiaorong Wang, Dexian Zheng, and Youhe Gao. A systematical analysis of tryptic peptide identification with reverse phase liquid chromatography and electrospray ion trap mass spectrometry. *Genomics, proteomics & bioinformatics*, 2(3):174–183, 2004.
- [87] Piero Giansanti, Liana Tsiatsiani, Teck Yew Low, and Albert J R Heck. Six alternative proteases for mass spectrometry–based proteomics beyond trypsin. *Nature Protocols*, 11(5):993–1006, 2016. doi: 10.1038/nprot.2016.057.

- [88] Elien Vandermarliere, Michael Mueller, and Lennart Martens. Getting intimate with trypsin, the leading protease in proteomics. *Mass Spectrometry Reviews*, 32 (6):453–465, nov 2013. ISSN 02777037. doi: 10.1002/mas.21376.
- [89] Promega. Proteases and surfactants, 2020. URL <https://nld.promega.com/products/mass-spectrometry/proteases-and-surfactants/>.
- [90] Lloyd R. Snyder, Joseph J. Kirkland, and John W. Dolan. *Introduction to Modern Liquid Chromatography*. John Wiley & Sons, Inc., Hoboken, NJ, USA, nov 2009. ISBN 9780470508183. doi: 10.1002/9780470508183.
- [91] Colin T Mant, Yuxin Chen, Zhe Yan, Traian V Popa, James M Kovacs, Janine B Mills, Brian P Tripet, and Robert S Hodges. HPLC Analysis and Purification of Peptides. In *Methods in Molecular Biology*, volume 386, chapter 1, pages 3–55. Springer, 2007.
- [92] P. Roepstorff and J. Fohlman. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biological Mass Spectrometry*, 11(11):601–601, nov 1984. ISSN 10969888. doi: 10.1002/bms.1200111109.
- [93] J.A McCloskey. *Mass spectrometry*. San Diego: Academic Press, 1990.
- [94] Leann M. Mikesch, Beatrix Ueberheide, An Chi, Joshua J. Coon, John E.P. Syka, Jeffrey Shabanowitz, and Donald F. Hunt. The utility of ETD mass spectrometry in proteomic analysis. *Biochimica et Biophysica Acta - Proteins and Proteomics*, 1764 (12):1811–1822, dec 2006. ISSN 15709639. doi: 10.1016/j.bbapap.2006.10.003.
- [95] Michaela Scigelova and Alexander Makarov. Orbitrap mass analyzer - Overview and applications in proteomics. In *Proteomics*, number 1-2 SUPPL., pages 16–21. John Wiley & Sons, Ltd, sep 2006. doi: 10.1002/pmic.200600528.
- [96] E de Hoffmann and V Stroobant. *Mass Spectrometry: Principles & Applications*. John Wiley & Sons, 3 edition, 2007.

- [97] Qizhi Hu, Robert J. Noll, Hongyan Li, Alexander Makarov, Mark Hardman, and R. Graham Cooks. The Orbitrap: a new mass spectrometer. *Journal of Mass Spectrometry*, 40(4):430–443, apr 2005. ISSN 1076-5174. doi: 10.1002/jms.856.
- [98] Roman A Zubarev and Alexander Makarov. Orbitrap mass spectrometry. *Analytical Chemistry*, 2013. doi: 10.1021/ac4001223.
- [99] Alan G. Marshall, Christopher L. Hendrickson, and George S. Jackson. Fourier transform ion cyclotron resonance mass spectrometry: A primer. *Mass Spectrometry Reviews*, 17(1):1–35, jan 1998. ISSN 1098-2787. doi: 10.1002/(SICI)1098-2787(1998)17:1<1::AID-MAS1>3.0.CO;2-K.
- [100] Michael P. Washburn, Dirk Wolters, and John R. Yates. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature Biotechnology*, 19(3):242–247, 2001. ISSN 10870156. doi: 10.1038/85686.
- [101] Jimmy K. Eng, Ashley L. McCormack, and John R. Yates. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *Journal of The American Society for Mass Spectrometry*, 5(11):976–989, 1994. doi: <https://doi.org/10.1021/jasms.8b00502>.
- [102] David N. Perkins, Darryl J. C. Pappin, David M. Creasy, and John S. Cottrell. Probability-based protein identification by searching sequence databases using mass spectrometry data. *ELECTROPHORESIS*, 20(18):3551–3567, dec 1999. ISSN 1522-2683. doi: 10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2.
- [103] Lewis Y. Geer, Sanford P. Markey, Jeffrey A. Kowalak, Lukas Wagner, Ming Xu, Dawn M. Maynard, Xiaoyu Yang, Wen Yao Shi, and Stephen H. Bryant. Open Mass Spectrometry Search Algorithm. *Journal of Proteome Research*, 3(5):958–64, 2004. doi: 10.1021/PR0499491.
- [104] R. Craig and R. C. Beavis. TANDEM: matching proteins with tandem mass spec-

- tra. *Bioinformatics*, 20(9):1466–1467, jun 2004. ISSN 1367-4803. doi: 10.1093/bioinformatics/bth092.
- [105] Jurgen Cox, Nadin Neuhauser, Annette Michalski, Richard A. Scheltema, Jesper V. Olsen, and Matthias Mann. Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. *Journal of Proteome Research*, 10(4): 1794–1805, apr 2011. ISSN 1535-3893. doi: 10.1021/pr101065j.
- [106] Benjamin J. Diament and William Stafford Noble. Faster SEQUEST Searching for Peptide Identification from Tandem Mass Spectra. *Journal of Proteome Research*, 10(9):3871–3879, sep 2011. ISSN 1535-3893. doi: 10.1021/pr101196n.
- [107] Jimmy K. Eng, Tahmina A. Jahan, and Michael R. Hoopmann. Comet: An open-source MS/MS sequence database search tool. *Proteomics*, 13(1):22–24, jan 2013. ISSN 16159853. doi: 10.1002/pmic.201200439.
- [108] Sangtae Kim and Pavel A. Pevzner. MS-GF+ makes progress towards a universal database search tool for proteomics. *Nature Communications*, 5(1):5277, dec 2014. ISSN 2041-1723. doi: 10.1038/ncomms6277.
- [109] Viktoria Dorfer, Peter Pichler, Thomas Stranzl, Johannes Stadlmann, Thomas Taus, Stephan Winkler, and Karl Mechtler. MS Amanda, a universal identification algorithm optimized for high accuracy tandem mass spectra. *Journal of Proteome Research*, 13(8):3679–3684, aug 2014. ISSN 15353907. doi: 10.1021/pr500202e.
- [110] Joao A Paulo. Practical and Efficient Searching in Proteomics: A Cross Engine Comparison. *WebmedCentral*, 4(10), oct 2013. ISSN 2046-1690. doi: 10.9754/journal.wplus.2013.0052.
- [111] Eugene A. Kapp, Frédéric Schütz, Lisa M. Connolly, John A. Chakel, Jose E. Meza, Christine A. Miller, David Fenyo, Jimmy K. Eng, Joshua N. Adkins, Gilbert S. Omenn, and Richard J. Simpson. An evaluation, comparison, and accurate benchmarking of several publicly available MS/MS search algorithms: Sensitiv-

- ity and specificity analysis. *Proteomics*, 5(13):3475–3490, aug 2005. ISSN 16159853. doi: 10.1002/pmic.200500126.
- [112] Yong Lin, Jian Zhou, Deng Bi, Ping Chen, Xianchun Wang, and Songping Liang. Sodium-deoxycholate-assisted tryptic digestion and identification of proteolytically resistant proteins. *Analytical Biochemistry*, 377(2):259–266, jun 2008. ISSN 0003-2697. doi: 10.1016/J.AB.2008.03.009.
- [113] Joshua E. Elias and Steven P. Gygi. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nature Methods*, 4(3):207–214, mar 2007. ISSN 15487091. doi: 10.1038/nmeth1019.
- [114] Suruchi Aggarwal and Amit Kumar Yadav. False discovery rate estimation in proteomics. In *Statistical Analysis in Proteomics*, pages 119–128. Springer, 2016.
- [115] John R. Yates, Scott F. Morgan, Christine L. Gatlin, Patrick R. Griffin, and Jimmy K. Eng. Method to Compare Collision-Induced Dissociation Spectra of Peptides: Potential for Library Searching and Subtractive Analysis. *Analytical Chemistry*, 70(17):3557–3565, sep 1998. ISSN 00032700. doi: 10.1021/ac980122y.
- [116] Henry Lam, Eric W. Deutsch, James S. Eddes, Jimmy K. Eng, Stephen E. Stein, and Ruedi Aebersold. Building consensus spectral libraries for peptide identification in proteomics. *Nature Methods*, 5(10):873–875, sep 2008. ISSN 15487091. doi: 10.1038/nmeth.1254.
- [117] Brendan MacLean, Daniela M. Tomazela, Nicholas Shulman, Matthew Chambers, Gregory L. Finney, Barbara Frewen, Randall Kern, David L. Tabb, Daniel C. Liebler, and Michael J. MacCoss. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics*, 26(7): 966–968, apr 2010. ISSN 1460-2059. doi: 10.1093/bioinformatics/btq054.
- [118] Chih-Chiang Tsou, Dmitry Avtonomov, Brett Larsen, Monika Tucholska, Hyungwon Choi, Anne-Claude Gingras, and Alexey I Nesvizhskii. DIA-Umpire:

- comprehensive computational framework for data-independent acquisition proteomics. *Nature Methods*, 12(3):258–264, mar 2015. ISSN 1548-7091. doi: 10.1038/nmeth.3255.
- [119] Hannes L Röst. Deep learning adds an extra dimension to peptide fragmentation. *Nature methods*, 16(6):469–470, 2019.
- [120] Yang-Ming Lin, Ching-Tai Chen, and Jia-Ming Chang. Ms2cnn: predicting ms/ms spectrum based on protein sequence using deep convolutional neural networks. *BMC genomics*, 20(9):1–10, 2019.
- [121] Shenheng Guan, Michael F Moran, and Bin Ma. Prediction of lc-ms/ms properties of peptides from sequence by deep learning. *Molecular & Cellular Proteomics*, 18(10):2099–2107, 2019.
- [122] Ngoc Hieu Tran, Xianglilan Zhang, Lei Xin, Baozhen Shan, and Ming Li. De novo peptide sequencing by deep learning. *Proceedings of the National Academy of Sciences*, 114(31):8247–8252, 2017.
- [123] Lichao Zhang and Joshua E. Elias. Relative protein quantification using tandem mass tag mass spectrometry. In *Methods in Molecular Biology*, volume 1550, pages 185–198. Humana Press Inc., 2017. doi: 10.1007/978-1-4939-6747-6\_14.
- [124] Thermo Fisher Scientific. Tandem Mass Tag (TMT) Systems, 2019.
- [125] Shao En Ong, Blagoy Blagoev, Irina Kratchmarova, Dan Bach Kristensen, Hanno Steen, Akhilesh Pandey, and Matthias Mann. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Molecular & cellular proteomics : MCP*, 1(5):376–386, may 2002. ISSN 15359476. doi: 10.1074/mcp.M200025-MCP200.
- [126] Matthias Mann. Functional and quantitative proteomics using SILAC. *Nature Reviews Molecular Cell Biology*, 7(12):952–958, dec 2006. ISSN 14710072. doi: 10.1038/nrm2067.

- [127] Philip L Ross, Yulin N Huang, Jason N Marchese, Brian Williamson, Kenneth Parker, Stephen Hattan, Nikita Khainovski, Sasi Pillai, Subhakar Dey, Scott Daniels, et al. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Molecular & cellular proteomics*, 3 (12):1154–1169, 2004.
- [128] Sebastian Wiese, Kai A. Reidegeld, Helmut E. Meyer, and Bettina Warscheid. Protein labeling by iTRAQ: A new tool for quantitative mass spectrometry in proteome research. *Proteomics*, 7(3):340–350, feb 2007. ISSN 16159853. doi: 10.1002/pmic.200600422.
- [129] Julia M. Burkhardt, Marc Vaudel, René P. Zahedi, Lennart Martens, and Albert Sickmann. iTRAQ protein quantification: A quality-controlled workflow. *PROTEOMICS*, 11(6):1125–1134, mar 2011. ISSN 16159853. doi: 10.1002/pmic.201000711.
- [130] Christopher M Shuford, James J Walters, Patricia M Holland, Uma Sreenivasan, Nadav Askari, Kevin Ray, and Russell P Grant. Absolute protein quantification by mass spectrometry: not as simple as advertised. *Analytical chemistry*, 89(14): 7406–7415, 2017.
- [131] Caroline Evans, Josselin Noirel, Saw Yen Ow, Malinda Salim, Ana G Pereira-Medrano, Narciso Couto, Jagroop Pandhal, Duncan Smith, Trong Khoa Pham, Esther Karunakaran, Xin Zou, Catherine A Biggs, and Wright Phillip C. An insight into iTRAQ: where do we stand now? *Analytical and bioanalytical chemistry*, 404(4):1011–1027, 2012.
- [132] Jürgen Cox, Marco Y. Hein, Christian A. Luber, Igor Paron, Nagarjuna Nagaraj, and Matthias Mann. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Molecular and Cellular Proteomics*, 13(9):2513–2526, sep 2014. ISSN 15359484. doi: 10.1074/mcp.M113.031591.

- [133] Melissa M. Matzke, Joseph N. Brown, Marina A. Gritsenko, Thomas O. Metz, Joel G. Pounds, Karin D. Rodland, Anil K. Shukla, Richard D. Smith, Katrina M. Waters, Jason E. McDermott, and Bobbie-Jo Webb-Robertson. A comparative analysis of computational approaches to relative protein quantification using peptide peak intensities in label-free LC-MS proteomics experiments. *PROTEOMICS*, 13(3-4):493–503, feb 2013. ISSN 16159853. doi: 10.1002/pmic.201200269.
- [134] Bobbie-Jo M. Webb-Robertson, Lee Ann McCue, Katrina M. Waters, Melissa M. Matzke, Jon M. Jacobs, Thomas O. Metz, Susan M. Varnum, and Joel G. Pounds. Combined Statistical Analyses of Peptide Intensities and Peptide Occurrences Improves Identification of Significant Peptides from MS-Based Proteomics Data. *Journal of Proteome Research*, 9(11):5748–5756, nov 2010. ISSN 1535-3893. doi: 10.1021/pr1005247.
- [135] Bobbie-Jo M. Webb-Robertson, Melissa M. Matzke, Jon M. Jacobs, Joel G. Pounds, and Katrina M. Waters. A statistical selection strategy for normalization procedures in LC-MS proteomics experiments through dataset-dependent ranking of normalization scaling factors. *PROTEOMICS*, 11(24):4736–4741, dec 2011. ISSN 16159853. doi: 10.1002/pmic.201100078.
- [136] Bobbie-Jo M. Webb-Robertson, Holli K. Wiberg, Melissa M. Matzke, Joseph N. Brown, Jing Wang, Jason E. McDermott, Richard D. Smith, Karin D. Rodland, Thomas O. Metz, Joel G. Pounds, and Katrina M. Waters. Review, Evaluation, and Discussion of the Challenges of Missing Value Imputation for Mass Spectrometry-Based Label-Free Global Proteomics. *Journal of Proteome Research*, 14(5):1993–2001, may 2015. ISSN 1535-3893. doi: 10.1021/pr501138h.
- [137] Allison Doerr. DIA mass spectrometry. *Nature Methods*, 12(1):35, jan 2014. ISSN 15487105. doi: 10.1038/nmeth.3234.
- [138] María Eugenia Monge and Facundo M. Fernández. An Introduction to Ambient Ionization Mass Spectrometry. In Marek Domin and Robert Cody, editors, *Am-*



- bient Ionization Mass Spectrometry*, chapter 1, pages 1–22. Royal Society of Chemistry, nov 2014. doi: 10.1039/9781782628026-00001.
- [139] Victor A. Mikhailov, Rian L. Griffiths, and Helen J. Cooper. Liquid extraction surface analysis for native mass spectrometry: Protein complexes and ligand binding. *International Journal of Mass Spectrometry*, 420:43–50, 2017. doi: 10.1016/j.ijms.2016.09.011.
- [140] RG Graham Graham Cooks, Zheng Ouyang, Zoltan Takats, and JM Justin M. Justin M Wiseman. Ambient Mass Spectrometry. *Science*, 311(5767), 2006.
- [141] Min Zong Huang, Sy Chi Cheng, Yi Tzu Cho, and Jentaie Shiea. Ambient ionization mass spectrometry: A tutorial. *Analytica Chimica Acta*, 702(1):1–15, sep 2011. ISSN 00032670. doi: 10.1016/j.aca.2011.06.017.
- [142] Zoltán Takáts, Justin M. Wiseman, and R. Graham Cooks. Ambient mass spectrometry using desorption electrospray ionization (DESI): instrumentation, mechanisms and applications in forensics, chemistry, and biology. *Journal of Mass Spectrometry*, 40(10):1261–1275, oct 2005. ISSN 1076-5174. doi: 10.1002/jms.922.
- [143] Livia S. Eberlin, Isaiah Norton, Daniel Orringer, Ian F. Dunn, Xiaohui Liu, Jennifer L. Ide, Alan K. Jarmusch, Keith L. Ligon, Ferenc A. Jolesz, Alexandra J. Golby, Sandro Santagata, Nathalie Y.R. Agar, and R. Graham Cooks. Ambient mass spectrometry for the intraoperative molecular diagnosis of human brain tumors. *Proceedings of the National Academy of Sciences of the United States of America*, 110(5):1611–1616, jan 2013. ISSN 00278424. doi: 10.1073/pnas.1215687110.
- [144] Zhong-Ping Yao. Characterization of proteins by ambient mass spectrometry. *Mass Spectrometry Reviews*, 31(4):437–447, jul 2012. ISSN 02777037. doi: 10.1002/mas.20346.
- [145] L.V. V Ratcliffe, F.J.M. J M Frank J. M. Rutten, D.A. A Barrett, T Whitmore, D Seymour, C Greenwood, Y Aranda-Gonzalvo, S Robinson, and M McCous-

- tra. Surface Analysis under Ambient Conditions Using Plasma-Assisted Desorption/Ionization Mass Spectrometry. *Analytical Chemistry*, 2007. doi: 10.1021/AC070109Q.
- [146] Daniel Eikel and Jack D. Henion. Liquid Extraction Surface Analysis Mass Spectrometry (LESA MS): Combining Liquid Extraction, Surface Profiling and Ambient Ionization Mass Spectrometry in One Novel Analysis Technique. In *Ambient Ionization Mass Spectrometry*, pages 482–495. Royal Society of Chemistry, 2014. doi: 10.1039/9781782628026-00482.
- [147] Rosana M. Alberici, Rosineide C. Simas, Gustavo B. Sanvido, Wanderson Romão, Priscila M. Lalli, Mario Benassi, Ildenize B. S. Cunha, and Marcos N. Eberlin. Ambient mass spectrometry: bringing MS into the “real world”. *Analytical and Bioanalytical Chemistry*, 398(1):265–294, sep 2010. ISSN 1618-2642. doi: 10.1007/s00216-010-3808-3.
- [148] M Montowska, MR R Alexander, GA A Tucker, and DA A Barrett. Authentication of processed meat products by peptidomic analysis using rapid ambient mass spectrometry. *Food chemistry*, 2015.
- [149] Magdalena Montowska, Morgan R. Alexander, Gregory A. Tucker, and David A. Barrett. Rapid Detection of Peptide Markers for Authentication Purposes in Raw and Cooked Meat Using Ambient Liquid Extraction Surface Analysis Mass Spectrometry. *Analytical Chemistry*, 86(20):10257–10265, oct 2014. ISSN 0003-2700. doi: 10.1021/ac502449w.
- [150] Magdalena Montowska, Wei Rao, Morgan R MR Alexander, Gregory A. Tucker, and David A. Barrett. Tryptic digestion coupled with ambient desorption electrospray ionization and liquid extraction surface analysis mass spectrometry enabling identification of skeletal muscle proteins in mixtures and distinguishing between beef, pork, horse, chicken, and t. *Analytical Chemistry*, 86(9):4479–4487, 2014.

- [151] Wei Rao, Adam D. Celiz, David J. Scurr, Morgan R. Alexander, and David A. Barrett. Ambient DESI and LESA-MS Analysis of Proteins Adsorbed to a Biomaterial Surface Using In-Situ Surface Tryptic Digestion. *Journal of The American Society for Mass Spectrometry*, 24(12):1927–1936, dec 2013. ISSN 1044-0305. doi: 10.1007/s13361-013-0737-3.
- [152] Z. Takats, Justin M Wiseman, Bogdan Gologan, and R Graham Cooks. Mass Spectrometry Sampling Under Ambient Conditions with Desorption Electrospray Ionization. *Science*, 306(5695):471–473, oct 2004. ISSN 0036-8075. doi: 10.1126/science.1104404.
- [153] Vilmos Kertesz and Gary J. Van Berkel. Fully automated liquid extraction-based surface sampling and ionization using a chip-based robotic nanoelectrospray platform. *Journal of Mass Spectrometry*, 45(3):252–260, mar 2010. ISSN 10765174. doi: 10.1002/jms.1709.
- [154] Klaudia I. Kocurek, Leanne Stones, Josephine Bunch, Robin C. May, and Helen J. Cooper. Top-Down LESA Mass Spectrometry Protein Analysis of Gram-Positive and Gram-Negative Bacteria. *Journal of The American Society for Mass Spectrometry*, pages 1–12, jul 2017. ISSN 1044-0305. doi: 10.1007/s13361-017-1718-8.
- [155] Elizabeth C. Randall, Josephine Bunch, and Helen J. Cooper. Direct analysis of intact proteins from *Escherichia coli* colonies by liquid extraction surface analysis mass spectrometry. *Analytical Chemistry*, 86(21):10504–10510, oct 2014. ISSN 15206882. doi: 10.1021/ac503349d.
- [156] Rebecca L. Edwards, Paul Griffiths, Josephine Bunch, and Helen J. Cooper. Top-Down Proteomics and Direct Surface Sampling of Neonatal Dried Blood Spots: Diagnosis of Unknown Hemoglobin Variants. *Journal of The American Society for Mass Spectrometry*, 23(11):1921–1930, nov 2012. ISSN 1044-0305. doi: 10.1007/s13361-012-0477-9.
- [157] Nicholas J. Martin, Josephine Bunch, and Helen J. Cooper. Dried Blood Spot Pro-

- teomics: Surface Extraction of Endogenous Proteins Coupled with Automated Sample Preparation and Mass Spectrometry Analysis. *Journal of The American Society for Mass Spectrometry*, 24(8):1242–1249, aug 2013. ISSN 1044-0305. doi: 10.1007/s13361-013-0658-1.
- [158] Nicholas J. Martin, Rian L. Griffiths, Rebecca L. Edwards, and Helen J. Cooper. Native Liquid Extraction Surface Analysis Mass Spectrometry: Analysis of Non-covalent Protein Complexes Directly from Dried Substrates. *Journal of The American Society for Mass Spectrometry*, 26(8):1320–1327, aug 2015. ISSN 1044-0305. doi: 10.1007/s13361-015-1152-8.
- [159] Rian L. Griffiths, Alex Dexter, Andrew J. Creese, and Helen J. Cooper. Liquid extraction surface analysis field asymmetric waveform ion mobility spectrometry mass spectrometry for the analysis of dried blood spots. *The Analyst*, 140(20): 6879–6885, 2015. ISSN 0003-2654. doi: 10.1039/C5AN00933B.
- [160] Rebecca L. Edwards, Andrew J. Creese, Mark Baumert, Paul Griffiths, Josephine Bunch, and Helen J. Cooper. Hemoglobin Variant Analysis via Direct Surface Sampling of Dried Blood Spots Coupled with High-Resolution Mass Spectrometry. *Analytical Chemistry*, 83(6):2265–2270, mar 2011. ISSN 0003-2700. doi: 10.1021/ac1030804.
- [161] Matthias Wilm and Matthias Mann. Analytical Properties of the Nanoelectrospray Ion Source. *Analytical Chemistry*, 68(1):1–8, jan 1996. ISSN 0003-2700. doi: 10.1021/ac9509519.
- [162] JB B Fenn, M Mann, CK K Meng, SF F Wong, and CM M Whitehouse. Electrospray ionization for mass spectrometry of large biomolecules. *Science*, 246(4926): 64–71, 1989.
- [163] John B Fenn. Electrospray Wings for Molecular Elephants (Nobel Lecture). *Angewandte Chemie International Edition*, 42(33):3871–3894, aug 2003. ISSN 1433-7851. doi: 10.1002/anie.200300605.

- [164] R Juraschek, T Dulcks, and M Karas. Nanoelectrospray - more than just a minimized-flow electrospray ionization source. *Journal of the American Society for Mass Spectrometry*, 10(4):300–308, apr 1999. ISSN 1044-0305. doi: 10.1016/S1044-0305(98)00157-3.
- [165] Richard B. Cole. *Electrospray and MALDI Mass Spectrometry: Fundamentals, Instrumentation, Practicalities, and Biological Applications*. Wiley & Sons, Hoboken, NJ, 2 edition, 2010. ISBN 978-0-471-74107-7.
- [166] Matthias Wilm. Principles of electrospray ionization. *Molecular and Cellular Proteomics*, 10(7):M111.009407, jul 2011. ISSN 15359476. doi: 10.1074/mcp.M111.009407.
- [167] J. V. Iribarne and B. A. Thomson. On the evaporation of small ions from charged droplets. *The Journal of Chemical Physics*, 64(6):2287–2294, aug 1976. ISSN 00219606. doi: 10.1063/1.432536.
- [168] Lord Rayleigh. On the equilibrium of liquid conducting masses charged with electricity. *Philosophical Magazine*, 14(87):184–186, 1882.
- [169] T.C. Rohner, N. Lion, and H.H. Girault. Electrochemical and theoretical aspects of electrospray ionisation. *Physical Chemistry Chemical Physics*, 6:3056–3068, 2004. doi: <https://doi.org/10.1039/B316836K>.
- [170] Paul D. Piehowski, Vladislav A. Petyuk, Daniel J. Orton, Fang Xie, Ronald J. Moore, Manuel Ramirez-Restrepo, Anzhelika Engel, Andrew P. Lieberman, Roger L. Albin, David G. Camp, Richard D. Smith, and Amanda J. Myers. Sources of technical variability in quantitative LC-MS proteomics: Human brain tissue sample analysis. *Journal of Proteome Research*, 12(5):2128–2137, may 2013. ISSN 15353907. doi: 10.1021/pr301146m.
- [171] Rian L. Griffiths and Helen J. Cooper. Direct Tissue Profiling of Protein Complexes: Toward Native Mass Spectrometry Imaging. *Analytical Chemistry*, 88(1): 606–609, jan 2016. ISSN 0003-2700. doi: 10.1021/acs.analchem.5b03993.

- [172] Rian L. Griffiths, Andrew J. Creese, Alan M. Race, Josephine Bunch, and Helen J. Cooper. LESA FAIMS Mass Spectrometry for the Spatial Profiling of Proteins from Tissue. *Analytical Chemistry*, 88(13):6758–6766, jul 2016. ISSN 0003-2700. doi: 10.1021/acs.analchem.6b01060.
- [173] Rian L. Griffiths, Elizabeth C. Randall, Alan M. Race, Josephine Bunch, and Helen J. Cooper. Raster-Mode Continuous-Flow Liquid Microjunction Mass Spectrometry Imaging of Proteins in Thin Tissue Sections. *Analytical Chemistry*, page acs.analchem.7b00977, may 2017. ISSN 0003-2700. doi: 10.1021/acs.analchem.7b00977.
- [174] Joscelyn Sarsby, RL L Rian L. Rian L Griffiths, AM Alan M. AM Alan M A M Race, Josephine Bunch, Elizabeth C. Randall, Andrew J. Creese, and Helen J. Cooper. Liquid extraction surface analysis mass spectrometry coupled with field asymmetric waveform ion mobility spectrometry for analysis of intact proteins from biological substrates. *Analytical Chemistry*, 87(13):6794–6800, jul 2015. ISSN 0003-2700. doi: 10.1021/acs.analchem.5b01151.
- [175] Rian L Griffiths, Albert Konijnenberg, Rosa Viner, and Helen J Cooper. Direct mass spectrometry analysis of protein complexes and intact proteins up to > 70 kda from tissue. *Analytical chemistry*, 91(11):6962–6966, 2019.
- [176] Emma K Sisley, Eva Illes-Toth, and Helen J Cooper. In situ analysis of intact proteins by ion mobility mass spectrometry. *TrAC Trends in Analytical Chemistry*, 124:115534, 2020.
- [177] Oliver J Hale, Emma K Sisley, Rian L Griffiths, Iain B Styles, and Helen J Cooper. Native lesa twims-msi: Spatial, conformational, and mass analysis of proteins and protein complexes. *Journal of the American Society for Mass Spectrometry*, 31(4):873–879, 2020.
- [178] Rian L Griffiths, Klaudia I Kocurek, and Helen J Cooper. Ambient surface mass

- spectrometry–ion mobility spectrometry of intact proteins. *Current Opinion in Chemical Biology*, 42:67–75, 2018.
- [179] Emma K Sisley, Jakub Ujma, Martin Palmer, Kevin Giles, Francisco A Fernandez-Lima, and Helen J Cooper. Lesa cyclic ion mobility mass spectrometry of intact proteins from thin tissue sections. *Analytical Chemistry*, 92(9):6321–6326, 2020.
- [180] Rian L. Griffiths, Anna L. Simmonds, John G. Swales, Richard J.A. Goodwin, and Helen J. Cooper. LESA MS Imaging of Heat-Preserved and Frozen Tissue: Benefits of Multistep Static FAIMS. *Analytical Chemistry*, 90(22):13306–13314, nov 2018. ISSN 15206882. doi: 10.1021/acs.analchem.8b02739.
- [181] Rian L. Griffiths, Emma K. Sisley, Andrea F. Lopez-Clavijo, Anna L. Simmonds, Iain B. Styles, and Helen J. Cooper. Native mass spectrometry imaging of intact proteins and protein complexes in thin tissue sections. *International Journal of Mass Spectrometry*, 437:23–29, mar 2019. ISSN 13873806. doi: 10.1016/j.ijms.2017.10.009.
- [182] Rian L. Griffiths, James W. Hughes, Susan E. Abbatiello, Michael W. Belford, Iain B. Styles, and Helen J. Cooper. Comprehensive LESA Mass Spectrometry Imaging of Intact Proteins by Integration of Cylindrical FAIMS. *Analytical Chemistry*, 92(4):2885–2890, feb 2020. ISSN 15206882. doi: 10.1021/acs.analchem.9b05124.
- [183] Ellen Scheerlinck, Maarten Dhaenens, Ann Van Soom, Luc Peelman, Petra De Sutter, Katleen Van Steendam, and Dieter Deforce. Minimizing technical variation during sample preparation prior to label-free quantitative mass spectrometry. *Analytical biochemistry*, 490:14–19, 2015.
- [184] Reinaldo Almeida, Zane Berzina, Eva C. Arnspang, Jan Baumgart, Johannes Vogt, Robert Nitsch, and Christer S. Ejsing. Quantitative Spatial Analysis of the Mouse Brain Lipidome by Pressurized Liquid Extraction Surface Ana-

- lysis. *Analytical Chemistry*, 87(3):1749–1756, feb 2015. ISSN 0003-2700. doi: 10.1021/ac503627z.
- [185] A. Gareth Brenton and A. Ruth Godfrey. Electro-Focusing Liquid Extractive Surface Analysis (EF-LESA) Coupled to Mass Spectrometry. *Analytical Chemistry*, 86(7):3323–3329, apr 2014. ISSN 0003-2700. doi: 10.1021/ac4035136.
- [186] Matthias Lorenz, Olga S. Ovchinnikova, and Gary J. Van Berkel. Fully automated laser ablation liquid capture surface analysis using nanoelectrospray ionization mass spectrometry. *Rapid Communications in Mass Spectrometry*, 28(11):1312–1320, jun 2014. ISSN 09514198. doi: 10.1002/rcm.6904.
- [187] Gary J. Van Berkel, Michael J. Ford, Mitchel J. Doktycz, and Stephen J. Kennel. Evaluation of a surface-sampling probe electrospray mass spectrometry system for the analysis of surface-deposited and affinity-captured proteins. *Rapid Communications in Mass Spectrometry*, 20(7):1144–1152, apr 2006. ISSN 0951-4198. doi: 10.1002/rcm.2428.
- [188] Anna A. Popova, Sebastian M. Schillo, Konstantin Demir, Erica Ueda, A. Nesterov-Mueller, and Pavel A. Levkin. Droplet-Array (DA) Sandwich Chip: A Versatile Platform for High-Throughput Cell Screening Based on Superhydrophobic-Superhydrophilic Micropatterning. *Advanced Materials*, 27(35):5217–5222, sep 2015. ISSN 09359648. doi: 10.1002/adma.201502115.
- [189] Florian L. Geyer, Erica Ueda, Urban Liebel, Nicole Grau, and Pavel A. Levkin. Superhydrophobic-Superhydrophilic Micropatterning: Towards Genome-on-a-Chip Cell Microarrays. *Angewandte Chemie International Edition*, 50(36):8424–8427, aug 2011. ISSN 14337851. doi: 10.1002/anie.201102545.
- [190] N. Valipour M., F.Ch. Birjandi, and J. Sargolzaei. Super-non-wettable surfaces: A review. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 448: 93–106, apr 2014. ISSN 0927-7757. doi: 10.1016/J.COLSURFA.2014.02.016.



- [191] Melanie J. Bailey, Elizabeth C. Randall, Catia Costa, Tara L. Salter, Alan M. Race, Marcel de Puit, Mattijs Koeberg, Mark Baumert, and Josephine Bunch. Analysis of urine, oral fluid and fingerprints by liquid extraction surface analysis coupled to high resolution MS and MS/MS – opportunities for forensic and biomedical science. *Anal. Methods*, 8(16):3373–3382, 2016. ISSN 1759-9660. doi: 10.1039/C6AY00782A.
- [192] Vilmos Kertesz and Gary J Van Berkel. Automated liquid microjunction surface sampling-HPLC-MS/MS analysis of drugs and metabolites in whole-body thin tissue sections. *Bioanalysis*, 5(7):819–826, apr 2013. ISSN 1757-6180. doi: 10.4155/bio.13.42.
- [193] Shane R. Ellis, Cameron J. Ferris, Kerry J. Gilmore, Todd W. Mitchell, Stephen J. Blanksby, and Marc in het Panhuis. Direct Lipid Profiling of Single Cells from Inkjet Printed Microarrays. *Analytical Chemistry*, 84(22):9679–9683, nov 2012. ISSN 0003-2700. doi: 10.1021/ac302634u.
- [194] Paul A. Moench, Alexandre Catoire, James Glick, and Jimmy Flarakos. Determination of tissue-specific ion suppression by liquid extraction surface analysis mass spectrometry. *Rapid Communications in Mass Spectrometry*, 30(2):340–342, jan 2016. ISSN 09514198. doi: 10.1002/rcm.7441.
- [195] Laura Tomlinson, Jens Fuchser, Arne Fütterer, Mark Baumert, David G. Hassall, Andrew West, and Peter S. Marshall. Using a single, high mass resolution mass spectrometry platform to investigate ion suppression effects observed during tissue imaging. *Rapid Communications in Mass Spectrometry*, 28(9):995–1003, may 2014. ISSN 09514198. doi: 10.1002/rcm.6869.
- [196] Daniel Eikel, Marissa Vavrek, Sheri Smith, Carol Bason, Suzie Yeh, Walter A. Korfmaier, and Jack D. Henion. Liquid extraction surface analysis mass spectrometry (LESA-MS) as a novel profiling tool for drug distribution and metabolism

- analysis: the terfenadine example. *Rapid Communications in Mass Spectrometry*, 25 (23):3587–3596, dec 2011. ISSN 09514198. doi: 10.1002/rcm.5274.
- [197] Z. Hall, Z. Ament, C. H. Wilson, D. L. Burkhardt, T. Ashmore, A. Koulman, T. Littlewood, G. I. Evan, and J. L. Griffin. Myc Expression Drives Aberrant Lipid Metabolism in Lung Cancer. *Cancer Research*, 76(16):4608–4618, aug 2016. ISSN 0008-5472. doi: 10.1158/0008-5472.CAN-15-3403.
- [198] Zoe Hall, Yajing Chu, and Julian L. Griffin. Liquid Extraction Surface Analysis Mass Spectrometry Method for Identifying the Presence and Severity of Non-alcoholic Fatty Liver Disease. *Analytical Chemistry*, 89(9):5161–5170, may 2017. ISSN 0003-2700. doi: 10.1021/acs.analchem.7b01097.
- [199] Darren Kessner, Matt Chambers, Robert Burke, David Agus, and Parag Mallick. ProteoWizard: open source software for rapid proteomics tools development. *Bioinformatics*, 24(21):2534–2536, nov 2008. ISSN 1460-2059. doi: 10.1093/bioinformatics/btn323.
- [200] Johannes Forkman. Estimator and Tests for Common Coefficients of Variation in Normal Distributions. *Communications in Statistics - Theory and Methods*, 38(2): 233–251, jan 2009. ISSN 0361-0926. doi: 10.1080/03610920802187448.
- [201] Alan M. Race, Andrew D. Palmer, Alex Dexter, Rory T. Steven, Iain B. Styles, and Josephine Bunch. SpectralAnalysis: Software for the Masses. *Analytical Chemistry*, 88(19):9451–9458, oct 2016. ISSN 0003-2700. doi: 10.1021/acs.analchem.6b01643.
- [202] A Savitzky and M.J.E Golay. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, 36(8):1627–1639, 1964. doi: 10.1021/ac60214a047.
- [203] Robert A van den Berg, Huub CJ Hoefsloot, Johan A Westerhuis, Age K Smilde, and Mariët J van der Werf. Centering, scaling, and transformations: improving

- the biological information content of metabolomics data. *BMC Genomics*, 7(1): 142, jun 2006. ISSN 14712164. doi: 10.1186/1471-2164-7-142.
- [204] Age K. Smilde, Mariët J. van der Werf, Sabina Bijlsma, Bianca J. C. van der Werff-van der Vat, and Renger H. Jellema. Fusion of Mass Spectrometry-Based Metabolomics Data. *Analytical Chemistry*, 77(20):6729–6736, oct 2005. ISSN 0003-2700. doi: 10.1021/ac051080y.
- [205] Stéphane Dray and Julie Josse. Principal component analysis with missing values: a comparative survey of methods. *Plant Ecology*, 216(5):657–667, may 2015. ISSN 1385-0237. doi: 10.1007/s11258-014-0406-z.
- [206] Yoav Benjamini and Yosef Hochberg. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1):289–300, 1995. doi: 10.2307/2346101.
- [207] C Hartmann, J Smeyers-Verbeke, D.L Massart, and R.D McDowall. Validation of bioanalytical chromatographic methods. *Journal of Pharmaceutical and Biomedical Analysis*, 17(2):193–218, jun 1998. ISSN 0731-7085. doi: 10.1016/S0731-7085(97)00198-2.
- [208] Bradley Worley and Robert Powers. Multivariate Analysis in Metabolomics. *Current Metabolomics*, 1(1):92–107, 2013.
- [209] David J. Clarke and Dominic J. Campopiano. Desalting large protein complexes during native electrospray mass spectrometry by addition of amino acids to the working solution. *The Analyst*, 140(8):2679–2686, mar 2015. ISSN 0003-2654. doi: 10.1039/C4AN02334J.
- [210] Sheng Pan, Ruedi Aebersold, Ru Chen, John Rush, David R Goodlett, Martin W McIntosh, Jing Zhang, and Teresa A Brentnall. Mass spectrometry based targeted protein quantification: methods and applications. *Journal of proteome research*, 8(2):787–797, 2009.

- [211] Joris Meurs, Morgan R. Alexander, Pavel A. Levkin, Simon Widmaier, Josephine Bunch, David A. Barrett, and Dong-Hyun Kim. Improved Extraction Repeatability and Spectral Reproducibility for Liquid Extraction Surface Analysis–Mass Spectrometry Using Superhydrophobic–Superhydrophilic Patterning. *Analytical Chemistry*, 90(10):6001–6005, may 2018. ISSN 0003-2700. doi: 10.1021/acs.analchem.8b00973.
- [212] Mark P. Molloy, Erin E. Brzezinski, Junqi Hang, Michael T. McDowell, and Ruth A. VanBogelen. Overcoming technical variation and biological variation in quantitative proteomics. *Proteomics*, 3(10):1912–1919, oct 2003. ISSN 16159853. doi: 10.1002/pmic.200300534.
- [213] Laetitia Cravello, David Lascoux, and Eric Forest. Use of different proteases working in acidic conditions to improve sequence coverage and resolution in hydrogen/deuterium exchange of large proteins. *Rapid communications in mass spectrometry*, 17(21):2387–2393, 2003.
- [214] Carolina Fernandez-Costa, Salvador Martínez-Bartolomé, Daniel B. McClatchy, Anthony J. Saviola, Nam-Kyung Yu, and John R. Yates. Impact of the identification strategy on the reproducibility of DDA and DIA results. *Journal of Proteome Research*, jun 2020. ISSN 1535-3893. doi: 10.1021/acs.jproteome.0c00153.
- [215] Léon Reubsaet, Michael J. Sweredoski, and Annie Moradian. Data-Independent Acquisition for the Orbitrap Q Exactive HF: A Tutorial. *Journal of Proteome Research*, 18(3):803–813, mar 2019. ISSN 1535-3893. doi: 10.1021/acs.jproteome.8b00845.
- [216] Ben C. Collins, Christie L. Hunter, Yansheng Liu, Birgit Schilling, George Rosenberger, Samuel L. Bader, Daniel W. Chan, Bradford W. Gibson, Anne Claude Gingras, Jason M. Held, Mio Hirayama-Kurogi, Guixue Hou, Christoph Krisp, Brett Larsen, Liang Lin, Siqi Liu, Mark P. Molloy, Robert L. Moritz, Sumio Ohtsuki, Ralph Schlapbach, Nathalie Selevsek, Stefani N. Thomas, Shin Cheng

- Tzeng, Hui Zhang, and Ruedi Aebersold. Multi-laboratory assessment of reproducibility, qualitative and quantitative performance of SWATH-mass spectrometry. *Nature Communications*, 8(1):1–12, dec 2017. ISSN 20411723. doi: 10.1038/s41467-017-00249-5.
- [217] Harald Barsnes and Marc Vaudel. SearchGUI: A Highly Adaptable Common Interface for Proteomics Search and de Novo Engines. *Journal of Proteome Research*, 17(7):2552–2555, jul 2018. ISSN 1535-3893. doi: 10.1021/acs.jproteome.8b00175.
- [218] Marc Vaudel, Julia M Burkhart, René P Zahedi, Eystein Oveland, Frode S Berven, Albert Sickmann, Lennart Martens, and Harald Barsnes. PeptideShaker enables reanalysis of MS-derived proteomics data sets. *Nature Biotechnology*, 33(1):22–24, jan 2015. ISSN 1087-0156. doi: 10.1038/nbt.3109.
- [219] David L. Tabb, Christopher G. Fernando, and Matthew C. Chambers. MyriMatch: Highly accurate tandem mass spectral peptide identification by multivariate hypergeometric analysis. *Journal of Proteome Research*, 6(2):654–661, feb 2007. ISSN 15353893. doi: 10.1021/pr0604054.
- [220] Jürgen Cox and Matthias Mann. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology*, 26(12):1367–1372, dec 2008. ISSN 1087-0156. doi: 10.1038/nbt.1511.
- [221] Shan-Shan Lin, Chi-Hong Wu, Mei-Chuan Sun, Chung-Ming Sun, and Yen-Peng Ho. Microwave-assisted enzyme-catalyzed reactions in various solvent systems. *Journal of the American Society for Mass Spectrometry*, 16(4):581–588, apr 2005. ISSN 1044-0305. doi: 10.1016/J.JASMS.2005.01.012.
- [222] Aimin Tan and John C. Fanaras. Use of high-pH (basic/alkaline) mobile phases for LC-MS or LC-MS/MS bioanalysis. *Biomedical Chromatography*, 33(1):e4409, jan 2019. ISSN 02693879. doi: 10.1002/bmc.4409.

- [223] M. R. Wilkins, E. Gasteiger, J. C. Sanchez, A. Bairoch, and D. F. Hochstrasser. Two-dimensional gel electrophoresis for proteome projects: The effects of protein hydrophobicity and copy number. *Electrophoresis*, 19(8-9):1501–1505, jun 1998. ISSN 01730835. doi: 10.1002/elps.1150190847.
- [224] Kai Zhang, Zhengjin Jiang, Chuanyi Yao, Zhichao Zhang, Qinsun Wang, Ruyu Gao, and Chao Yan. Separation of peptides by pressurized capillary electrochromatography. *Journal of Chromatography A*, 987(1-2):453–458, feb 2003. ISSN 00219673. doi: 10.1016/S0021-9673(02)02063-0.
- [225] Ambrose Furey, Merisa Moriarty, Vaishali Bane, Brian Kinsella, and Mary Lehan. Ion suppression; A critical review on causes, evaluation, prevention and applications, oct 2013. ISSN 00399140.
- [226] Joerg Doellinger, Marica Grossegeesse, Andreas Nitsche, and Peter Lasch. DMSO as a mobile phase additive enhances detection of ubiquitination sites by nano-LC-ESI-MS/MS. *Journal of Mass Spectrometry*, 53(2):183–187, feb 2018. ISSN 10765174. doi: 10.1002/jms.4049.
- [227] Hannes Hahne, Fiona Pachl, Benjamin Ruprecht, Stefan K. Maier, Susan Klaeger, Dominic Helm, Guillaume Médard, Matthias Wilm, Simone Lemeer, and Bernhard Kuster. DMSO enhances electrospray response, boosting sensitivity of proteomic experiments. *Nature Methods*, 10(10):989–991, oct 2013. ISSN 15487091. doi: 10.1038/nmeth.2610.
- [228] Dominika Strzelecka, Stephen W Holman, and Claire E Evers. Evaluation of dimethyl sulfoxide (DMSO) as a mobile phase additive during top 3 label-free quantitative proteomics. *International journal of mass spectrometry*, 391:157–160, nov 2015. doi: 10.1016/j.ijms.2015.07.004.
- [229] Péter Judák, Janelle Grainger, Catrin Goebel, Peter Van Eenoo, and Koen Deventer. DMSO Assisted Electrospray Ionization for the Detection of Small Peptide Hormones in Urine by Dilute-and-Shoot-Liquid-Chromatography-High

- Resolution Mass Spectrometry. *Journal of the American Society for Mass Spectrometry*, 28(8):1657–1665, aug 2017. ISSN 18791123. doi: 10.1007/s13361-017-1670-7.
- [230] Magdalena Montowska and Anita Spychaj. Quantification of species-specific meat proteins in cooked and smoked sausages using infusion mass spectrometry. *Journal of Food Science and Technology*, 55(12):4984–4993, dec 2018. ISSN 09758402. doi: 10.1007/s13197-018-3437-y.
- [231] Tommi Välikangas, Tomi Suomi, and Laura L. Elo. A systematic evaluation of normalization methods in quantitative label-free proteomics. *Briefings in Bioinformatics*, 19(1):1–11, 2018. doi: <https://doi.org/10.1093/bib/bbw095>.
- [232] Thomas M Annesley. Ion suppression in mass spectrometry. *Clinical chemistry*, 49(7):1041–4, jul 2003. ISSN 0009-9147. doi: 10.1373/49.7.1041.
- [233] Unige A Laskay, Anna A Lobas, Kristina Srzentic, Mikhail V Gorshkov, and Yury O Tsybin. Proteome digestion specificity analysis for rational design of extended bottom-up and middle-down proteomics experiments. *Journal of proteome research*, 12(12):5558–5569, 2013.
- [234] Tsutomu Arakawa, Yoshiko Kita, and Serge N. Timasheff. Protein precipitation and denaturation by dimethyl sulfoxide. *Biophysical Chemistry*, 131(1-3):62–70, 2007.
- [235] D. H. Rammner. The effect of DMSO on several enzyme systems. *Annals of the New York Academy of Sciences*, 141(1 Biological Ac):291–299, mar 1967. ISSN 0077-8923. doi: 10.1111/j.1749-6632.1967.tb34893.x.
- [236] Daniel J. Ryan, Nathan Heath Patterson, Nicole E. Putnam, Aimee D. Wilde, Andy Weiss, William J. Perry, James E. Cassat, Eric P. Skaar, Richard M. Caprioli, and Jeffrey M. Spraggins. MicroLESA: Integrating Autofluorescence Microscopy, In Situ Micro-Digestions, and Liquid Extraction Surface Analysis for

- High Spatial Resolution Targeted Proteomic Studies. *Analytical Chemistry*, 91(12): 7578–7585, jun 2019. ISSN 0003-2700. doi: 10.1021/acs.analchem.8b05889.
- [237] Lindsay K. Pino, Brian C. Searle, Han-Yin Yang, Andrew N. Hoofnagle, William S. Noble, and Michael J. MacCoss. Matrix-matched calibration curves for assessing analytical figures of merit in quantitative proteomics. *Journal of Proteome Research*, page acs.jproteome.9b00666, feb 2020. ISSN 1535-3893. doi: 10.1021/acs.jproteome.9b00666.
- [238] Jocelyn F. Krey, Phillip A. Wilmarth, Jung-Bum Shin, John Klimek, Nicholas E. Sherman, Erin D. Jeffery, Dongseok Choi, Larry L. David, and Peter G. Barr-Gillespie. Accurate Label-Free Protein Quantitation with High- and Low-Resolution Mass Spectrometers. *Journal of Proteome Research*, 13(2):1034–1044, feb 2014. ISSN 1535-3893. doi: 10.1021/pr401017h.
- [239] Melissa M. Matzke, Katrina M. Waters, Thomas O. Metz, Jon M. Jacobs, Amy C. Sims, Ralph S. Baric, Joel G. Pounds, and Bobbie-Jo M. Webb-Robertson. Improved quality control processing of peptide-centric LC-MS proteomics data. *Bioinformatics*, 27(20):2866–2872, oct 2011. ISSN 1460-2059. doi: 10.1093/bioinformatics/btr479.
- [240] Yuliya V Karpievitch, Alan R Dabney, and Richard D Smith. Normalization and missing value imputation for label-free LC-MS analysis. *BMC Bioinformatics*, 13 (S16):S5, nov 2012. ISSN 1471-2105. doi: 10.1186/1471-2105-13-S16-S5.
- [241] Jakob Willforss, Aakash Chawade, and Fredrik Levander. NormalyzerDE: On-line Tool for Improved Normalization of Omics Expression Data and High-Sensitivity Differential Expression Analysis. *Journal of Proteome Research*, 18(2): 732–740, feb 2019. ISSN 1535-3893. doi: 10.1021/acs.jproteome.8b00523.
- [242] Marcus Bantscheff, Markus Schirle, Gavain Sweetman, Jens Rick, and Bernhard Kuster. Quantitative mass spectrometry in proteomics: A critical review. *Analyt-*



- ical and Bioanalytical Chemistry*, 389(4):1017–1031, oct 2007. ISSN 16182642. doi: 10.1007/s00216-007-1486-6.
- [243] Adam D. Celiz, James G. W. Smith, Robert Langer, Daniel G. Anderson, David A. Winkler, David A. Barrett, Martyn C. Davies, Lorraine E. Young, Chris Denning, and Morgan R. Alexander. Materials for stem cell factories of the future. *Nature Materials*, 13(6):570–579, may 2014. ISSN 1476-1122. doi: 10.1038/nmat3972.
- [244] James G.W. GW W Smith, Adam D. Celiz, Asha K. Patel, Robert D. Short, Morgan R. Alexander, and Chris Denning. Scaling human pluripotent stem cell expansion and differentiation: are cell factories becoming a reality? *Regenerative Medicine*, 10(8):925–930, nov 2015. ISSN 1746-0751. doi: 10.2217/rme.15.65.
- [245] Rong Zhang, Heidi K. Mjoseng, Marieke A. Hoeve, Nina G. Bauer, Steve Pells, Rut Besseling, Srinivas Velugotla, Guilhem Tourniaire, Ria E.B. Kishen, Yanina Tsenkina, Chris Armit, Cairnan R.E. Duffy, Martina Helfen, Frank Edenhofer, Paul A. De Sousa, and Mark Bradley. A thermoresponsive and chemically defined hydrogel for long-term culture of human embryonic stem cells. *Nature Communications*, 4(1):1–10, jan 2013. ISSN 20411723. doi: 10.1038/ncomms2341.
- [246] Chengde Gao, Shuping Peng, Pei Feng, and Cijun Shuai. Bone biomaterials and interactions with stem cells. *Bone Research*, 5(1):1–33, dec 2017. ISSN 20956231. doi: 10.1038/boneres.2017.59.
- [247] Peptrotech. Biogems, 2020. URL <https://www.peptrotech.com/gb/>.
- [248] Robin Thompson. A note on restricted maximum likelihood estimation with an alternative outlier model. *Journal of the Royal Statistical Society: Series B (Methodological)*, 47(1):53–55, 1985.
- [249] Cosmin Lazar, Laurent Gatto, Myriam Ferro, Christophe Bruley, and Thomas Burger. Accounting for the Multiple Natures of Missing Values in Label-Free Quantitative Proteomics Data Sets to Compare Imputation Strategies. *Journal of*

- Proteome Research*, 15(4):1116–1125, apr 2016. ISSN 1535-3893. doi: 10.1021/acs.jproteome.5b00981.
- [250] Johan A. Westerhuis, Huub C.J. Hoefsloot, Suzanne Smit, Daniel J. Vis, Age K. Smilde, Ewoud J.J. Velzen, John P.M. Duijnhoven, and Ferdi A. Dorsten. Assessment of PLSDA cross validation. *Metabolomics*, 4(1):81–89, mar 2008. ISSN 15733882. doi: 10.1007/s11306-007-0099-6.
- [251] Sarah Krainer, Chris Smit, and Ulrich Hirn. The effect of viscosity and surface tension on inkjet printed picoliter dots. *RSC Advances*, 9(54):31708–31719, oct 2019. ISSN 20462069. doi: 10.1039/c9ra04993b.
- [252] Mark K. McQuain, Kevin Seale, Joel Peek, Shawn Levy, and Frederick R. Haselton. Effects of relative humidity and buffer additives on the contact printing of microarrays by quill pins. *Analytical Biochemistry*, 320(2):281–291, sep 2003. ISSN 00032697. doi: 10.1016/S0003-2697(03)00348-8.
- [253] Lianjiang Tan, Huifang Chen, Ding Pan, and Ning Pan. Investigation into the gelation and crystallization of polyacrylonitrile. *European Polymer Journal*, 45(5):1617–1624, may 2009. ISSN 00143057. doi: 10.1016/j.eurpolymj.2008.12.006.
- [254] Andrew L Hook and David J Scurr. ToF-sims analysis of a polymer microarray composed of poly (meth) acrylates with c6 derivative pendant groups. *Surface and Interface Analysis*, 48(4):226–236, 2016.
- [255] Roxanne H. Croze, William J. Thi, and Dennis O. Clegg. ROCK inhibition promotes attachment, proliferation, and wound closure in human embryonic stem cell-derived retinal pigmented epithelium. *Translational Vision Science and Technology*, 5(6), nov 2016. ISSN 21642591. doi: 10.1167/tvst.5.6.7.
- [256] Jinku Kim. Systematic approach to characterize the dynamics of protein adsorption on the surface of biomaterials using proteomics. *Colloids and Surfaces B: Biointerfaces*, 188:110756, 2020.

- [257] Ziryan Othman, Berta Cillero Pastor, Sabine van Rijt, and Pamela Habibovic. Understanding interactions between biomaterials and biological systems using proteomics. *Biomaterials*, 167:191–204, 2018.
- [258] Hassan M. Rostam, Leanne E. Fisher, Andrew L. Hook, Laurence Burroughs, Jeni C. Lockett, Graziela P. Figueredo, Chidimma Mbadugha, Alvin C.K. Teo, Arsalan Latif, Lisa Kämmerling, Mitchell Day, Karen Lawler, David Barrett, So-maia Elsheikh, Mohammad Ilyas, David A. Winkler, Morgan R. Alexander, and Amir M. Ghaemmaghani. Immune-Instructive Polymers Control Macrophage Phenotype and Modulate the Foreign Body Response In Vivo. *Matter*, 2(6):1564–1581, jun 2020. ISSN 25902385. doi: 10.1016/j.matt.2020.03.018.
- [259] Guokai Chen, Daniel R Gulbranson, Zhonggang Hou, Jennifer M Bolin, Victor Ruotti, Mitchell D Probasco, Kimberly Smuga-Otto, Sara E Howden, Nicole R Diol, Nicholas E Propson, Ryan Wagner, Garrett O Lee, Jessica Antosiewicz-Bourget, Joyce M C Teng, and James A Thomson. Chemically defined conditions for human iPSC derivation and culture. *Nature Methods*, 8(5):424–429, may 2011. ISSN 1548-7091. doi: 10.1038/nmeth.1593.
- [260] Na Young Ha, Shin Hye Kim, Tae Geol Lee, and Sang Yun Han. Rapid Characterization of Protein Chips Using Microwave-Assisted Protein Tryptic Digestion and MALDI Mass Spectrometry. *Langmuir*, 27(16):10098–10105, aug 2011. ISSN 0743-7463. doi: 10.1021/la201812a.
- [261] Frank E. Grubbs. Sample Criteria for Testing Outlying Observations. *The Annals of Mathematical Statistics*, 21(1):27–58, mar 1950. ISSN 0003-4851. doi: 10.1214/aoms/1177729885.
- [262] Olga Troyanskaya, Michael Cantor, Gavin Sherlock, Pat Brown, Trevor Hastie, Robert Tibshirani, David Botstein, and Russ B. Altman. Missing value estimation methods for DNA microarrays. *Bioinformatics*, 17(6):520–525, jun 2001. ISSN 13674803. doi: 10.1093/bioinformatics/17.6.520.

- [263] Vincent Sarrazy, Anne Koehler, Melissa L. Chow, Elena Zimina, Chen X. Li, Hideyuki Kato, Christopher A. Caldarone, and Boris Hinz. Integrins  $\alpha v\beta 5$  and  $\alpha v\beta 3$  promote latent TGF- $\beta 1$  activation by human cardiac fibroblast contraction. *Cardiovascular Research*, 102(3):407–417, jun 2014. ISSN 17553245. doi: 10.1093/cvr/cvu053.
- [264] Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, and Peter Walter. Integrins. In *Molecular Biology of the Cell*. Garland Science, 4 edition, 2002.
- [265] John G. Swales, Nicole Strittmatter, James W. Tucker, Malcolm R. Clench, Peter J. H. Webborn, and Richard J. A. Goodwin. Spatial Quantitation of Drugs in tissues using Liquid Extraction Surface Analysis Mass Spectrometry Imaging. *Scientific Reports*, 6(1):37648, dec 2016. ISSN 2045-2322. doi: 10.1038/srep37648.
- [266] Jana Havlikova, Elizabeth C. Randall, Rian L. Griffiths, John G. Swales, Richard J.A. Goodwin, Josephine Bunch, Iain B. Styles, and Helen J. Cooper. Quantitative Imaging of Proteins in Tissue by Stable Isotope Labeled Mimetic Liquid Extraction Surface Analysis Mass Spectrometry. *Analytical Chemistry*, 2019. ISSN 15206882. doi: 10.1021/acs.analchem.9b04148.
- [267] Anna M Kotowska, Gustavo F Trindade, Paula M Mendes, Philip M Williams, Jonathan W Aylott, Alexander G Shard, Morgan R Alexander, and David J Scurr. Protein identification by 3d orbisims to facilitate in situ imaging and depth profiling. *Nature communications*, 11(1):1–8, 2020.
- [268] Malin Andersson, M. Reid Groseclose, Ariel Y. Deutch, and Richard M. Caprioli. Imaging mass spectrometry of proteins and peptides: 3D volume reconstruction. *Nature Methods*, 5(1):101–108, jan 2008. ISSN 15487091. doi: 10.1038/nmeth1145.
- [269] Richard M. Caprioli, Terry B. Farmer, and Jocelyn Gile. Molecular Imaging of Biological Samples: Localization of Peptides and Proteins Using MALDI-TOF

- MS. *Analytical Chemistry*, 69(23):4751–4760, dec 1997. ISSN 00032700. doi: 10.1021/ac970888i.
- [270] Liam A. McDonnell, Garry L. Corthals, Stefan M. Willems, Alexandra van Reemoortere, René J.M. van Zeijl, and André M. Deelder. Peptide and protein imaging mass spectrometry in cancer research. *Journal of Proteomics*, 73(10):1921–1944, sep 2010. ISSN 18743919. doi: 10.1016/j.jprot.2010.05.007.
- [271] Richard J.A. Goodwin, Stephen R. Pennington, and Andrew R. Pitt. Protein and peptides in pictures: Imaging with MALDI mass spectrometry, sep 2008. ISSN 16159853.
- [272] Christopher D. Bennett, Sarah E. Kohe, Simrandip K. Gill, Nigel P. Davies, Martin Wilson, Lisa C. D. Storer, Timothy Ritzmann, Simon M. L. Paine, Ian S. Scott, Ina Nicklaus-Wollenteit, Daniel A. Tennant, Richard G. Grundy, and Andrew C. Peet. Tissue metabolite profiles for the characterisation of paediatric cerebellar tumours. *Scientific Reports*, 8(1):11992, dec 2018. ISSN 2045-2322. doi: 10.1038/s41598-018-30342-8.
- [273] John-Paul Kilday, Ruman Rahman, Sara Dyer, Lee Ridley, James Lowe, Beth Coyle, and Richard Grundy. Pediatric Ependymoma: Biological Perspectives. *Molecular Cancer Research*, 7(6):765–786, jun 2009. ISSN 1541-7786. doi: 10.1158/1541-7786.MCR-08-0584.
- [274] Jing Wu, Terry S. Armstrong, and Mark R. Gilbert. Biology and management of ependymomas. *Neuro-Oncology*, 18(7):902–913, 2016. doi: <https://doi.org/10.1093/neuonc/now016>.
- [275] B. Messahel, S. Ashley, F. Saran, D. Ellison, J. Ironside, K. Phipps, T. Cox, W. K. Chong, K. Robinson, S. Picton, C. R. Pinkerton, C. Mallucci, D. Macarthur, T. Jaspan, A. Michalski, and R. G. Grundy. Relapsed intracranial ependymoma in children in the UK: Patterns of relapse, survival and therapeutic out-

- come. *European Journal of Cancer*, 45(10):1815–1823, jul 2009. ISSN 09598049. doi: 10.1016/j.ejca.2009.03.018.
- [276] Paul Miggiels, Bert Wouters, Gerard J.P. van Westen, Anne-Charlotte Dubbelman, and Thomas Hankemeier. Novel technologies for metabolomics: More for less. *TrAC Trends in Analytical Chemistry*, nov 2018. ISSN 0165-9936. doi: 10.1016/J.TRAC.2018.11.021.
- [277] Arun Sreekumar, Laila M. Poisson, Thekkelnaycke M. Rajendiran, Amjad P. Khan, Qi Cao, Jindan Yu, Bharathi Laxman, Rohit Mehra, Robert J. Lonigro, Yong Li, Mukesh K. Nyati, Aarif Ahsan, Shanker Kalyana-Sundaram, Bo Han, Xuhong Cao, Jaeman Byun, Gilbert S. Omenn, Debashis Ghosh, Subramaniam Pennathur, Danny C. Alexander, Alvin Berger, Jeffrey R. Shuster, John T. Wei, Sooryanarayana Varambally, Christopher Beecher, and Arul M. Chinnaiyan. Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. *Nature*, 457(7231):910–914, feb 2009. ISSN 0028-0836. doi: 10.1038/nature07762.
- [278] Q. Huang, Y. Tan, P. Yin, G. Ye, P. Gao, X. Lu, H. Wang, and G. Xu. Metabolic Characterization of Hepatocellular Carcinoma Using Nontargeted Tissue Metabolomics. *Cancer Research*, 73(16):4992–5002, aug 2013. ISSN 0008-5472. doi: 10.1158/0008-5472.CAN-13-0308.
- [279] Richard Beger. A Review of Applications of Metabolomics in Cancer. *Metabolites*, 3(3):552–574, jul 2013. ISSN 2218-1989. doi: 10.3390/metabo3030552.
- [280] Renu Pandey, Laura Caflisch, Alessia Lodi, Andrew J. Brenner, and Stefano Tiziani. Metabolomic signature of brain cancer. *Molecular Carcinogenesis*, 56(11):2355–2371, nov 2017. doi: 10.1002/mc.22694.
- [281] Caroline H. Johnson, Julijana Ivanisevic, and Gary Siuzdak. Metabolomics: beyond biomarkers and towards mechanisms. *Nature Reviews Molecular Cell Biology*, 17(7):451–459, jul 2016. ISSN 1471-0072. doi: 10.1038/nrm.2016.25.

- [282] Juha Kononen, Lukas Bubendorf, Anne Kallionimeni, Maarit Bärklund, Peter Schraml, Stephen Leighton, Joachim Torhorst, Michael J Mihatsch, Guido Sauter, and Olli-P. Kallionimeni. Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nature Medicine*, 4(7):844–847, jul 1998. ISSN 1078-8956. doi: 10.1038/nm0798-844.
- [283] M.A. Rubin, Rodney Dunn, Myla Strawderman, and K.J. Pienta. Tissue Microarray Sampling Strategy for Prostate Cancer Biomarker Analysis. *The American Journal of Surgical Pathology*, 26(3):312–319, 2002.
- [284] H Battifora. The Multitumor (Sausage) Tissue Block: Novel Method for Immunohistochemical Antibody Testing. *Laboratory Investigation*, 55(2):244–248, 1986.
- [285] Sankha S. Basu, Elizabeth C. Randall, Michael S. Regan, Begoña G.C. Lopez, Amanda R. Clark, Nicholas D. Schmitt, Jeffrey N. Agar, Deborah A. Dillon, and Nathalie Y.R. Agar. In Vitro Liquid Extraction Surface Analysis Mass Spectrometry (ivLESA-MS) for Direct Metabolic Analysis of Adherent Cells in Culture. *Analytical Chemistry*, 90(8):4987–4991, apr 2018. ISSN 15206882. doi: 10.1021/acs.analchem.8b00530.
- [286] Diren Beyoğlu and Jeffrey R. Idle. Metabolomics and its potential in drug development. *Biochemical Pharmacology*, 85(1):12–20, jan 2013. ISSN 18732968. doi: 10.1016/j.bcp.2012.08.013.
- [287] Mukesh Verma, Muin J. Khoury, and John P.A. Ioannidis. *Opportunities and challenges for selected emerging technologies in cancer epidemiology: Mitochondrial, epigenomic, metabolomic, and telomerase profiling*, volume 22. American Association for Cancer Research, feb 2013. doi: 10.1158/1055-9965.EPI-12-1263.
- [288] Alice Ly, Achim Buck, Benjamin Balluff, Na Sun, Karin Gorzolka, Annette Feuchtinger, Klaus-Peter Janssen, Peter J K Kuppen, Cornelis J H van de Velde, Gregor Weirich, Franziska Erlmeier, Rupert Langer, Michaela Aubele, Horst

- Zitzelsberger, Liam McDonnell, Michaela Aichler, and Axel Walch. High-mass-resolution MALDI mass spectrometry imaging of metabolites from formalin-fixed paraffin-embedded tissue. *Nature Protocols*, 11(8):1428–1443, aug 2016. ISSN 1754-2189. doi: 10.1038/nprot.2016.081.
- [289] Riccardo Di Guida, Jasper Engel, J. William Allwood, Ralf J. M. Weber, Martin R. Jones, Ulf Sommer, Mark R. Viant, and Warwick B. Dunn. Non-targeted UHPLC-MS metabolomic data processing methods: a comparative investigation of normalisation, missing value imputation, transformation and scaling. *Metabolomics*, 12(5):93, may 2016. ISSN 1573-3882. doi: 10.1007/s11306-016-1030-9.
- [290] Valentina Pirro, Paolo Oliveri, Christina Ramires Ferreira, Andrés Felipe González-Serrano, Zoltan Machaty, and Robert Graham Cooks. Lipid characterization of individual porcine oocytes by dual mode DESI-MS and data fusion. *Analytica Chimica Acta*, 848:51–60, oct 2014. ISSN 18734324. doi: 10.1016/j.aca.2014.08.001.
- [291] David S. Wishart, Yannick Djoumbou Feunang, Ana Marcu, An Chi Guo, Kevin Liang, Rosa Vázquez-Fresno, Tanvir Sajed, Daniel Johnson, Carin Li, Naama Karu, Zinat Sayeeda, Elvis Lo, Nazanin Assempour, Mark Berjanskii, Sandeep Singhal, David Arndt, Yonjie Liang, Hasan Badran, Jason Grant, Arnau Serra-Cayuela, Yifeng Liu, Rupa Mandal, Vanessa Neveu, Allison Pon, Craig Knox, Michael Wilson, Claudine Manach, and Augustin Scalbert. HMDB 4.0: The human metabolome database for 2018. *Nucleic Acids Research*, 46(D1):D608–D617, jan 2018. ISSN 13624962. doi: 10.1093/nar/gkx1089.
- [292] L Cottret, D Wildridge, F Vinson, M.P Barrett, H Charles, M.-F Sagot, and F Jourdan. MetExplore: a web server to link metabolomic experiments and genome-scale metabolic networks. *Nucleic Acids Research*, 38(2):W132–W137, 2010. doi: <https://doi.org/10.1093/nar/gkq312>.
- [293] L Cottret, C Frainay, M Chazalviel, F Cabanettes, Y Gloaguen, E Camenen, B Mer-



- let, S Heux, J.-C Portais, N Poupin, F Vinson, and F Jourdan. MetExplore: collaborative edition and exploration of metabolic networks. *Nucleic Acids Research*, 46 (W1):W495–W502, 2018. doi: <https://doi.org/10.1093/nar/gky301>.
- [294] Alan M. Race, Iain B. Styles, and Josephine Bunch. Inclusive sharing of mass spectrometry imaging data requires a converter for all. *Journal of Proteomics*, 75 (16):5111–5112, aug 2012. ISSN 1874-3919. doi: 10.1016/J.JPROT.2012.05.035.
- [295] Corbin G. Thompson, Mark T. Bokhart, Craig Sykes, Lourdes Adamson, Yuri Fedoriw, Paul A. Luciw, David C. Muddiman, Angela D.M. Kashuba, and Elias P. Rosen. Mass spectrometry imaging reveals heterogeneous efavirenz distribution within putative HIV reservoirs. *Antimicrobial Agents and Chemotherapy*, 59 (5):2944–2948, may 2015. ISSN 10986596. doi: 10.1128/AAC.04952-14.
- [296] Matthieu Peyre, Frédéric Commo, Carmela Dantas-Barbosa, Felipe Andreiuolo, Stéphanie Puget, Ludovic Lacroix, Françoise Drusch, Véronique Scott, Pascale Varlet, Audrey Mauguén, Philippe Dessen, Vladimir Lazar, Gilles Vassal, and Jacques Grill. Portrait of ependymoma recurrence in children: Biomarkers of tumor progression identified by dual-color microarray-based gene expression analysis. *PLoS ONE*, 5(9), 2010. ISSN 19326203. doi: 10.1371/journal.pone.0012932.
- [297] Lindsey M. Hoffman, Andrew M. Donson, Ichiro Nakachi, Andrea M. Griesinger, Diane K. Birks, Vladimir Amani, Molly S. Hemenway, Arthur K. Liu, Michael Wang, Todd C. Hankinson, Michael H. Handler, and Nicholas K. Foreman. Molecular sub-group-specific immunophenotypic changes are associated with outcome in recurrent posterior fossa ependymoma. *Acta Neuropathologica*, 127(5): 731–745, nov 2014. ISSN 14320533. doi: 10.1007/s00401-013-1212-8.
- [298] Gordon K. Smyth and Terry Speed. Normalization of cDNA microarray data. *Methods*, 31(4):265–273, 2003. ISSN 10462023. doi: 10.1016/S1046-2023(03)00155-5.
- [299] W. Evan Johnson, Cheng Li, and Ariel Rabinovic. Adjusting batch effects in mi-

- croarray expression data using empirical Bayes methods. *Biostatistics*, 8(1):118–127, jan 2007. ISSN 14654644. doi: 10.1093/biostatistics/kxj037.
- [300] Jelle J. Goeman, Sara Van de Geer, Floor De Kort, and Hans C. van Houwelingen. A global test for groups of genes: Testing association with a clinical outcome. *Bioinformatics*, 20(1):93–99, jan 2004. ISSN 13674803. doi: 10.1093/bioinformatics/btg382.
- [301] S. Cuellar-Baena, J.M. Morales, H. Martinetto, J. Calvar, G. Sevillever, G. Castellano, M. Cerda-Nicolas, B. Celda, and D. Monleon. Comparative metabolic profiling of paediatric ependymoma, medulloblastoma and pilocytic astrocytoma. *International Journal of Molecular Medicine*, 26(6):941–948, oct 2010. ISSN 11073756. doi: 10.3892/ijmm\_00000546.
- [302] Allen E Waziri, Michael Wang, Nicholas K Foreman Handler, Bette K Kleinschmidt-DeMasters, Michael H Andrew M Donson, Diane K Birks, and Valerie N Barton. Ependymoma Associated with a Good Prognosis in Immune Gene and Cell Enrichment Is. *J Immunol References*, 183:7428–7440, 2009. doi: 10.4049/jimmunol.0902811.
- [303] Andrew M. Donson, Diane K. Birks, Valerie N. Barton, Qi Wei, Bette K. Kleinschmidt-DeMasters, Michael H. Handler, Allen E. Waziri, Michael Wang, and Nicholas K. Foreman. Immune Gene and Cell Enrichment Is Associated with a Good Prognosis in Ependymoma. *The Journal of Immunology*, 183(11):7428–7440, dec 2009. ISSN 0022-1767. doi: 10.4049/jimmunol.0902811.
- [304] S. John Liu, Stephen T. Magill, Harish N. Vasudevan, Stephanie Hilz, Javier E. Villanueva-Meyer, Sydney Lastella, Vikas Daggubati, Jordan Spatz, Abrar Choudhury, Brent A. Orr, Benjamin Demaree, Kyounghee Seo, Sean P. Ferris, Adam R. Abate, Nancy Ann Oberheim Bush, Andrew W. Bollen, Michael W. McDermott, Joseph F. Costello, and David R. Raleigh. Multiplatform Molecular Profiling Reveals Epigenomic Intratumor Heterogeneity in Ependymoma. *Cell*

*Reports*, 30(5):1300–1309.e5, feb 2020. ISSN 22111247. doi: 10.1016/j.celrep.2020.01.018.

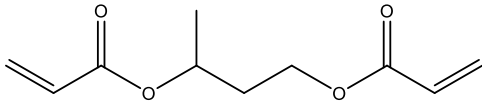
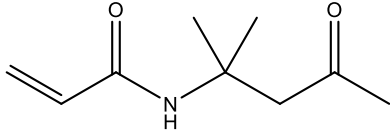
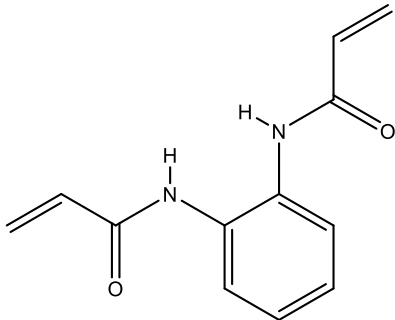
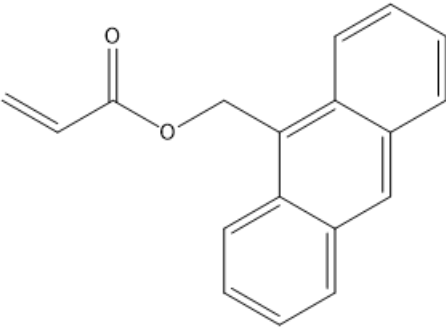
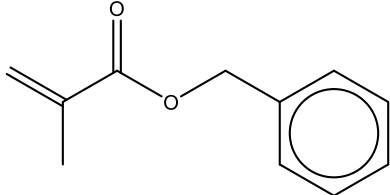
- [305] Kulandaimanuvel Antony Michealraj, Sachin A. Kumar, Leo J.Y. Kim, Florence M.G. Cavalli, David Przelicki, John B. Wojcik, Alberto Delaidelli, Andrea Bajic, Olivier Saulnier, Graham MacLeod, Ravi N. Vellanki, Maria C. Vladoiu, Paul Guilhamon, Winnie Ong, John J.Y. Lee, Yanqing Jiang, Borja L. Holgado, Alex Rasnitsyn, Ahmad A. Malik, Ricky Tsai, Cory M. Richman, Kyle Juraschka, Joonas Haapasalo, Evan Y. Wang, Pasqualino De Antonellis, Hiromichi Suzuki, Hamza Farooq, Polina Balin, Kaitlin Kharas, Randy Van Ommeren, Olga Sirbu, Avesta Rastan, Stacey L. Krumholtz, Michelle Ly, Mouloud Ahmadi, Geneviève Deblois, Dilakshan Srikanthan, Betty Luu, James Loukides, Xiaochong Wu, Livia Garzia, Vijay Ramaswamy, Evgeny Kanshin, María Sánchez-Osuna, Ibrahim El-Hamamy, Fiona J. Coutinho, Panagiotis Prianos, Sheila Singh, Laura K. Donovan, Craig Daniels, Daniel Schramek, Mike Tyers, Samuel Weiss, Lincoln D. Stein, Mathieu Lupien, Bradley G. Wouters, Benjamin A. Garcia, Cheryl H. Arrowsmith, Poul H. Sorensen, Stephane Angers, Nada Jabado, Peter B. Dirks, Stephen C. Mack, Sameer Agnihotri, Jeremy N. Rich, and Michael D. Taylor. Metabolic Regulation of the Epigenome Drives Lethal Infantile Ependymoma. *Cell*, 181(6), jun 2020. ISSN 10974172. doi: 10.1016/j.cell.2020.04.047.
- [306] Kristian W. Pajtler, Hendrik Witt, Martin Sill, David T.W. Jones, Volker Hovestadt, Fabian Kratochwil, Khalida Wani, Ruth Tatevossian, Chandanamali Punchihewa, Pascal Johann, Jüri Reimand, Hans Jörg Warnatz, Marina Ryzhova, Steve Mack, Vijay Ramaswamy, David Capper, Leonille Schweizer, Laura Sieber, Andrea Wittmann, Zhiqin Huang, Peter van Sluis, Richard Volckmann, Jan Koster, Rogier Versteeg, Daniel Fults, Helen Toledano, Smadar Avigad, Lindsey M. Hoffman, Andrew M. Donson, Nicholas Foreman, Ekkehard Hewer, Karel Zitterbart, Mark Gilbert, Terri S. Armstrong, Nalin Gupta, Jeffrey C. Al-

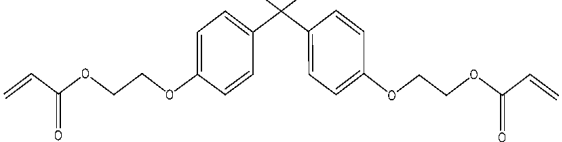
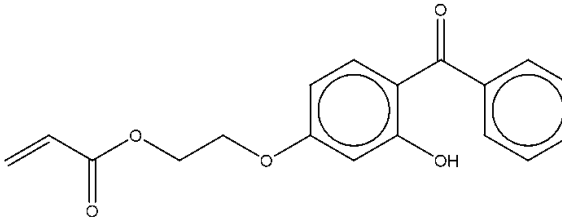
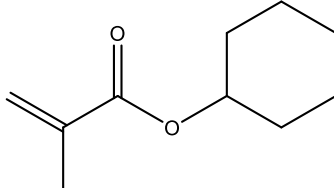
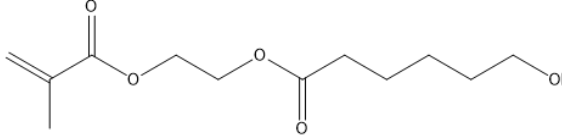
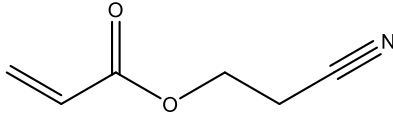
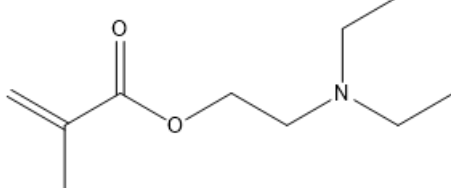
- len, Matthias A. Karajannis, David Zagzag, Martin Hasselblatt, Andreas E. Kulozik, Olaf Witt, V. Peter Collins, Katja von Hoff, Stefan Rutkowski, Torsten Pietsch, Gary Bader, Marie Laure Yaspo, Andreas von Deimling, Peter Lichter, Michael D. Taylor, Richard Gilbertson, David W. Ellison, Kenneth Aldape, Andrey Korshunov, Marcel Kool, and Stefan M. Pfister. Molecular Classification of Ependymal Tumors across All CNS Compartments, Histopathological Grades, and Age Groups. *Cancer Cell*, 2015. ISSN 18783686. doi: 10.1016/j.ccell.2015.04.002.
- [307] B. Balluff, M. Hanselmann, and R. M.A. Heeren. Mass Spectrometry Imaging for the Investigation of Intratumor Heterogeneity. In *Advances in Cancer Research*, volume 134, pages 201–230. Academic Press Inc., jan 2017. ISBN 9780128052495. doi: 10.1016/bs.acr.2016.11.008.
- [308] M.A. Qazi, P. Vora, C. Venugopal, S.S. Sidhu, J. Moffat, C. Swanton, and S.K. Singh. Intratumoral heterogeneity: pathways to treatment resistance and relapse in human glioblastoma. *Annals of Oncology*, 28(7):1448–1456, 2017. doi: <https://doi.org/10.1093/annonc/mdx169>.
- [309] Santiago Ramón y Cajal, Marta Sesé, Claudia Capdevila, Trond Aasen, Leticia De Mattos-Arruda, Salvador J. Diaz-Cano, Javier Hernández-Losa, and Josep Castellví. Clinical implications of intratumor heterogeneity: challenges and opportunities. *Journal of Molecular Medicine*, 98(2):161–177, feb 2020. ISSN 14321440. doi: 10.1007/s00109-020-01874-2.
- [310] C. Allison Stewart, Carl M. Gay, Yuanxin Xi, Santhosh Sivajothi, V. Sivakamasundari, Junya Fujimoto, Mohan Bolisetty, Patrice M. Hartsfield, Veerakumar Balasubramaniyan, Milind D. Chalise, Cesar Moran, Neda Kalhor, John Stewart, Hai Tran, Stephen G. Swisher, Jack A. Roth, Jianjun Zhang, John de Groot, Bonnie Glisson, Trudy G. Oliver, John V. Heymach, Ignacio Wistuba, Paul Robson, Jing Wang, and Lauren Averett Byers. Single-cell analyses reveal increased intratumoral heterogeneity after the onset of therapy

resistance in small-cell lung cancer. *Nature Cancer*, 1(4):423–436, apr 2020. doi:  
10.1038/s43018-019-0020-z.

## APPENDIX A

# **Acronyms and structural formulas for used monomers**

Acronym	Name	Structure
13BDDA	Butanediol-1,3 diacrylate	
AAcAm	Diacetone acrylamide	
AcAPAm	N-[2-(Acryloylamino)phenyl]acrylamide	
AnMA	Anthracenylmethacrylate	
BnMA	Benzyl methacrylate	

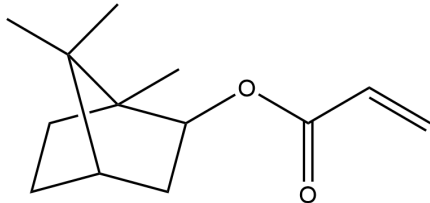
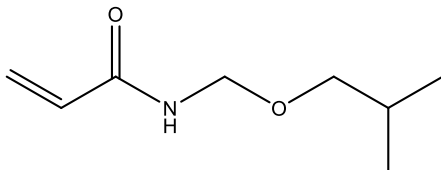
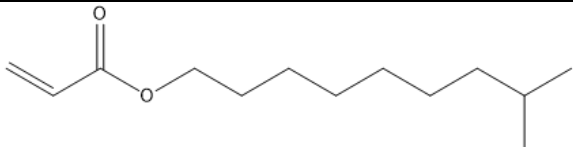
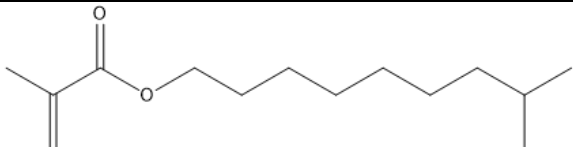
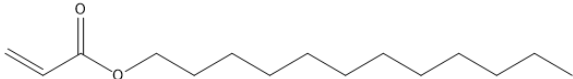
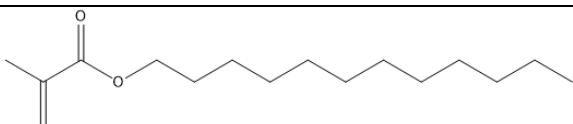
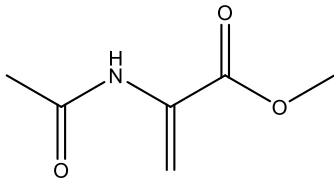
Acronym	Name	Structure
BPEODA	Bisphenol A ethoxylate diacrylate	
BzHPEA	Benzoyl-3-hydroxy-phenoxy)ethyl acrylate	
CHMA	Cyclohexyl methacrylate	
CMAOE	Caprolactone 2-(methacryloyloxy)ethyl ester	
CNEA	Cyanoethyl acrylate	
DEAEMA	Diethylaminoethyl methacrylate	

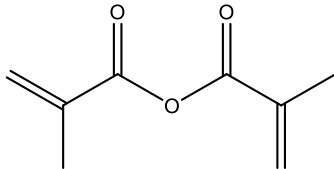
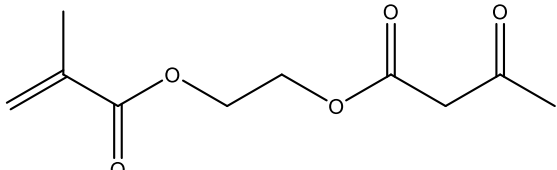
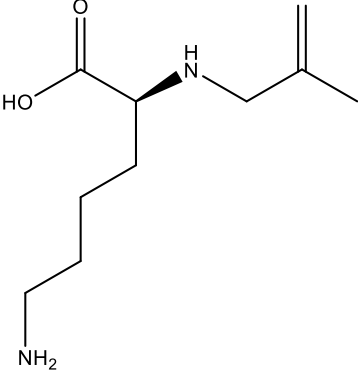
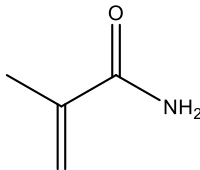
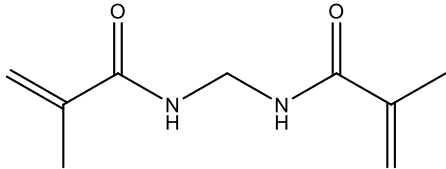


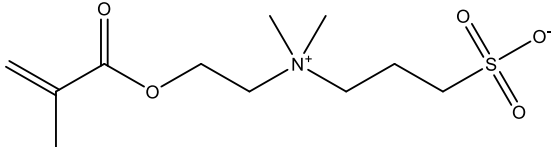
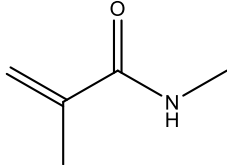
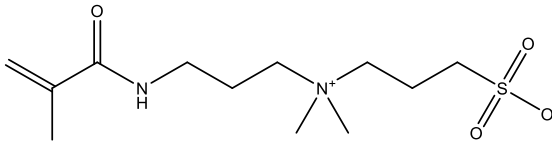
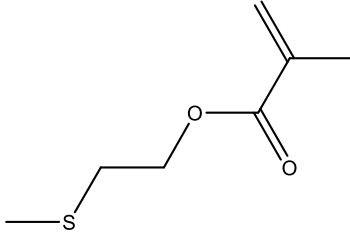
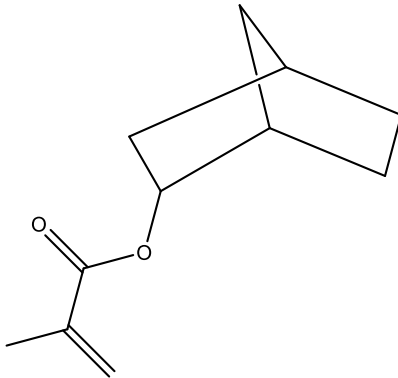
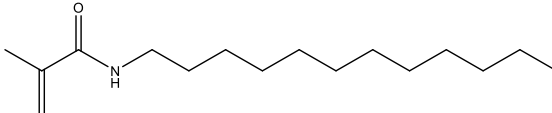
Acronym	Name	Structure
DEGDMA	Diethylene glycol dimethacrylate	
DEGEEA	Di(ethylene glycol) ethyl ether acrylate	
DEGEHA	Di(ethylene glycol) 2-ethylhexyl ether acrylate	
EaNIA	Ethyl trans-α-cyano-3-indole-acrylate	
EbCNA	Ethyl-cis-B-cyano-acrylate	
EG4DMA	Tetraethylene glycol dimethacrylate	
EGDPEA	Ethylene glycol dicyclopentenyl ether acrylate	

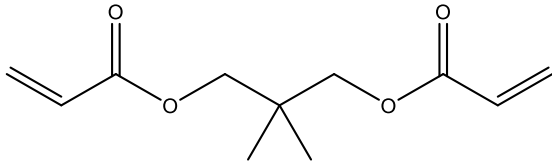
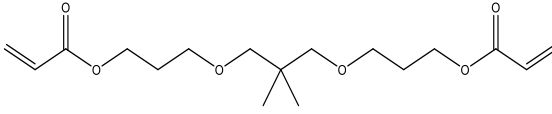
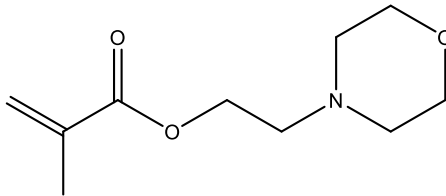
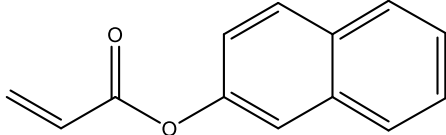
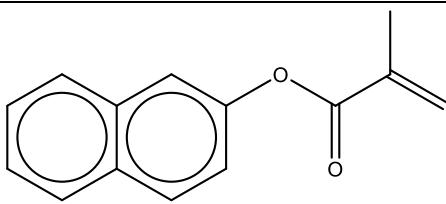
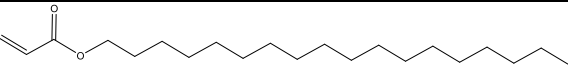
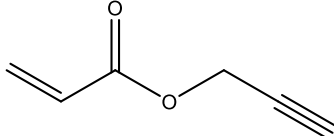
Acronym	Name	Structure
EGDMA	Ethylene glycol dimethacrylate	
EGMEA	Ethylene glycol methyl ether acrylate	
EGPEA	Ethylene glycol phenyl ether acrylate	
EGPhMA	Ethylene glycol phenyl ether methacrylate	
EOEA	Ethoxyethyl acrylate	
FMHPNMA	Trifluoro-2'-(trifluoromethyl)-2'-hydroxypropyl]-3-norbornyl methacrylate	

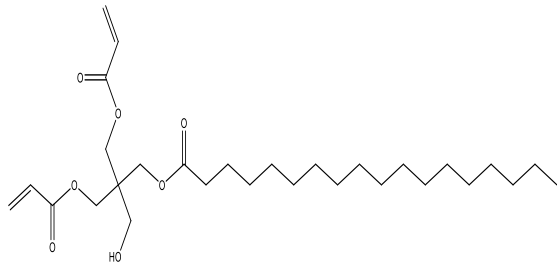
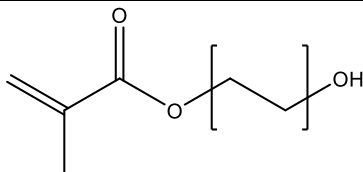
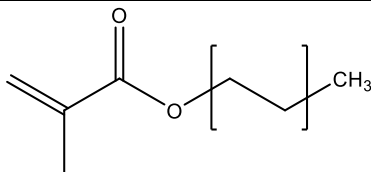
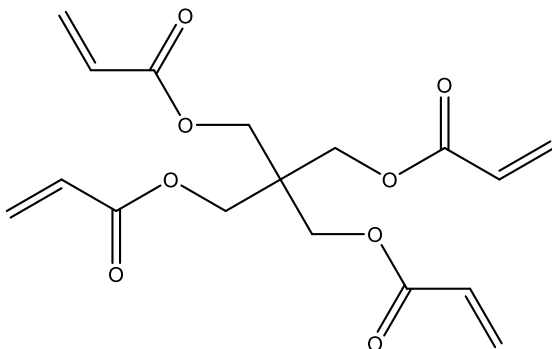
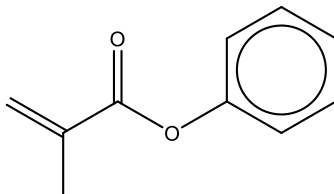
Acronym	Name	Structure
HDDMA	1,6-Hexanediol dimethacrylate,	
HFiPA	Hexafluoroisopropyl acrylate	
HFiPMA	Hexafluoroisopropyl methacrylate	
HMA	Hexyl methacrylate	
HPA	Hydroxypropyl acrylate	
iBA	Isobutyl acrylate	
iBMA	Isobornyl methacrylate	

Acronym	Name	Structure
iBOA	Isobornyl acrylate	
iBOMAm	N-(Isobutoxymethyl)acrylamide	
iDA	Isodecyl acrylate	
iDMA	Isodecyl methacrylate	
LaA	Lauryl acrylate	
LMA	Lauryl methacrylate	
MAA	Methyl 2-acetamidoacrylate	

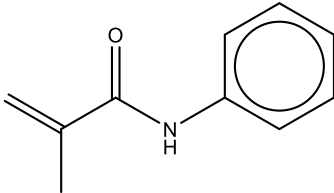
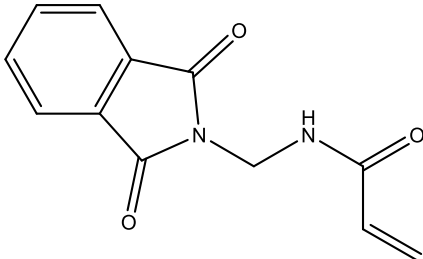
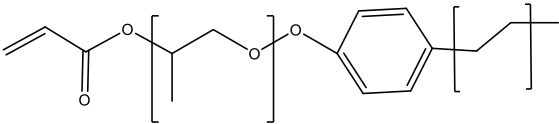
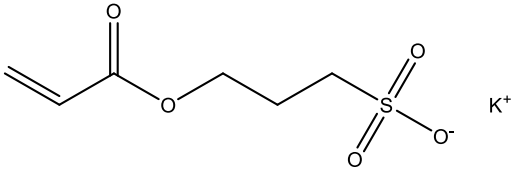
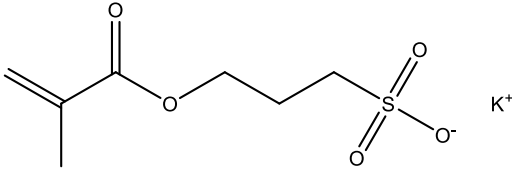
Acronym	Name	Structure
MAAH	Methacrylic anhydride	
MAEA	Methacryloyloxy)ethyl acetoacetate	
MAL	Methacryloyl-L-Lysine	
Mam	Methacrylamide	
MBMAm	N,N'-Methylenebismethacrylamide	

Acronym	Name	Structure
MEDMSAH	[2-(Methacryloyloxy)ethyl]dimethyl-(3-sulfopropyl) ammonium hydroxide	
MMAm	N-Methylmethacrylamide	
MPDSAH	Methacryloylamino)propyl]dimethyl(3-sulfopropyl)ammonium hydroxide inner salt	
MTEMA	Methylthioethyl methacrylate	
NBMA	Norbornyl methacrylate	
NDMAm	N-Dodecylmethacrylamide	

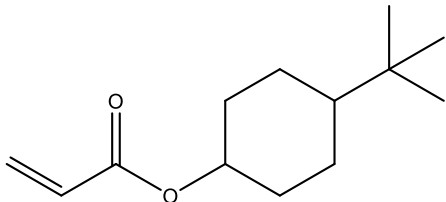
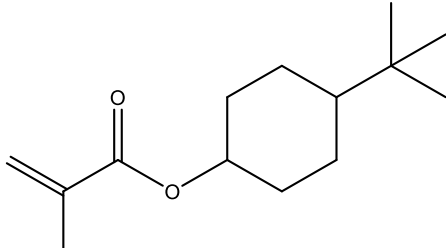
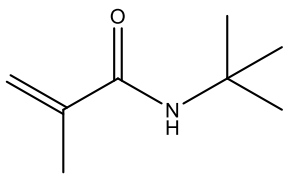
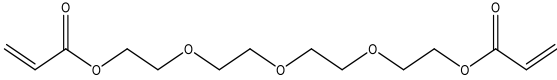
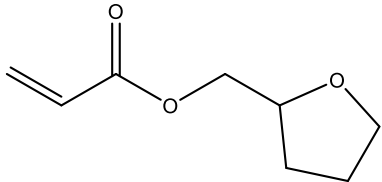
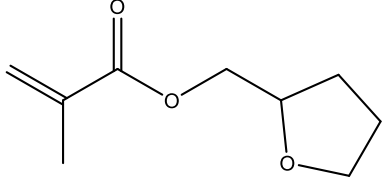
Acronym	Name	Structure
NGDA	Neopentyl glycol diacrylate	
NGPDA	Neopentyl glycol propoxylate diacrylate	
NMEMA	2-N-Morpholinoethyl methacrylate	
NpA	Naphthyl acrylate	
NpMA	Naphthyl methacrylate	
ODA	Octadecyl acrylate	
PA	Propargyl acrylate	

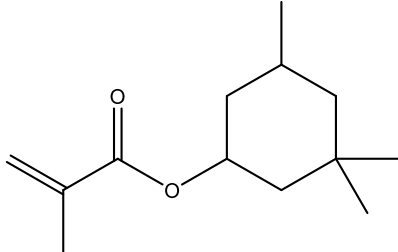
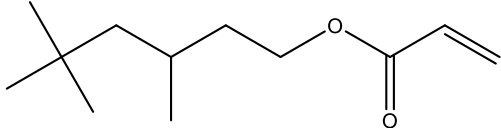
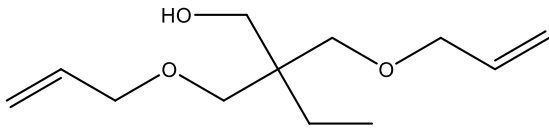
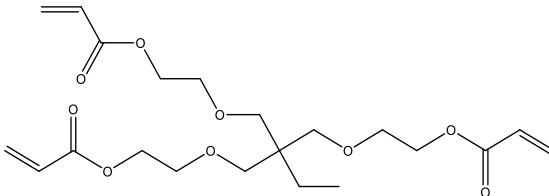
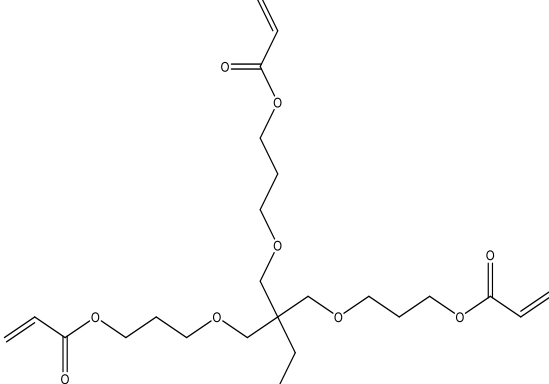
Acronym	Name	Structure
PEDAM	Pentaerythritol diacrylate monostearate	
pEGMA	Poly(ethylene glycol) methacrylate	
pEGMEMA	Poly(ethylene glycol) methyl ether methacrylate	
PETA	Pentaerythritol tetraacrylate	
PhMA	Phenyl methacrylate	

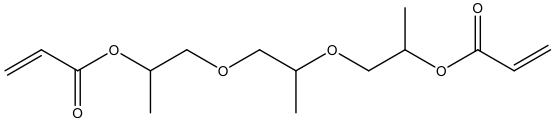
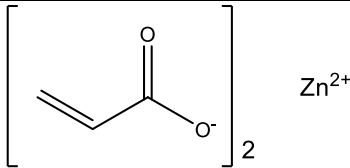
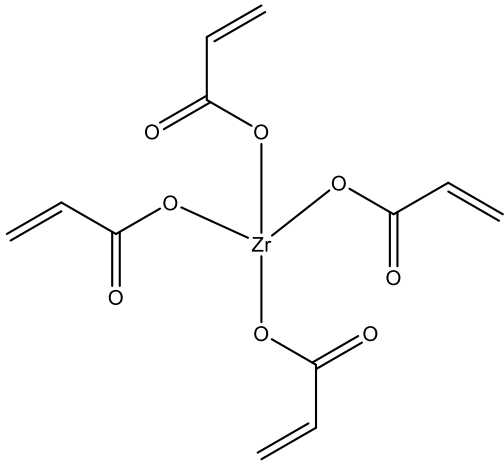
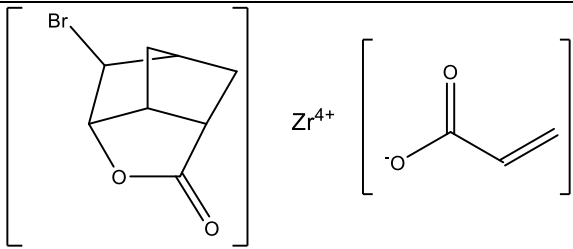
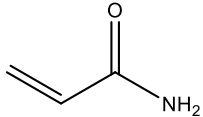


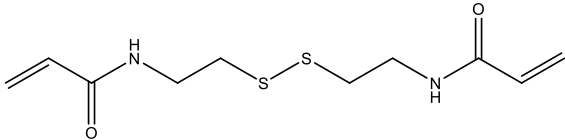
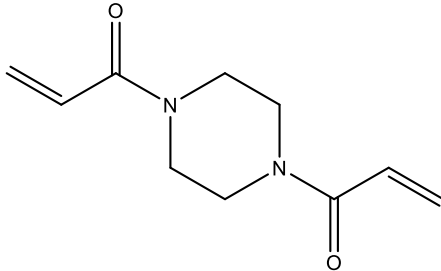
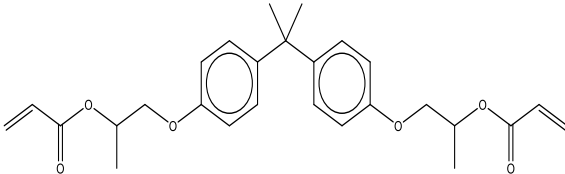
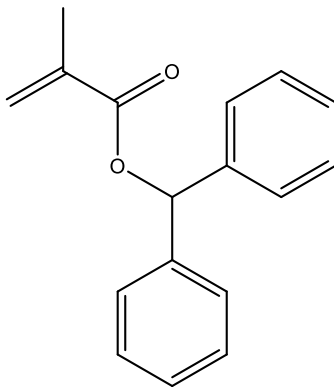
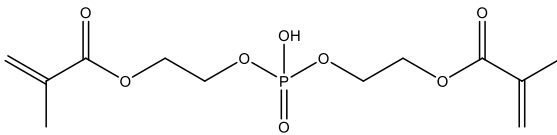
Acronym	Name	Structure
PhMAm	N-Phenylmethacrylamide	
PMAm	N-(Phthalimidomethyl)acrylamide	
pPGNEA	Poly(propylene glycol) 4-nonylphenyl ether acrylate	
SPAK	Sulfopropyl acrylate potassium salt	
SPMAK	3-Sulfopropyl methacrylate potassium salt	

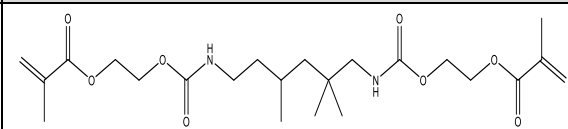
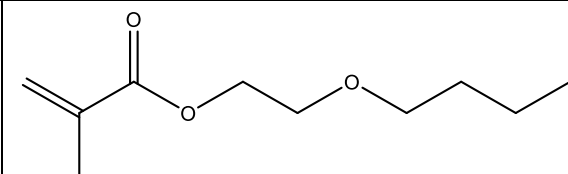
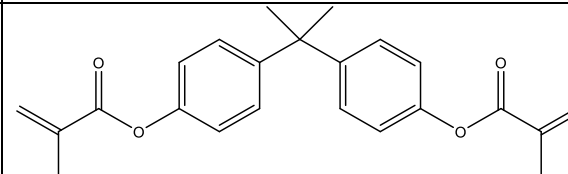
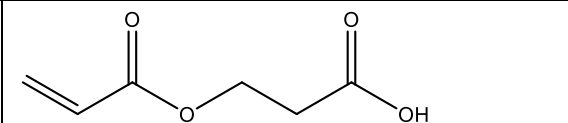
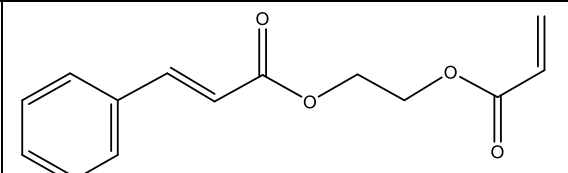
Acronym	Name	Structure
TAHTA	1,3,5-Triacryloylhexahydro-1,3,5-triazine	
TAIC	Tris[2-(acryloyloxy)ethyl] isocyanurate	
tBA	Tert-butyl acrylate	
tBAEMA	Tert-butylamino-ethyl methacrylate	
tBA <sub>m</sub>	N-tert-Butylacrylamide	

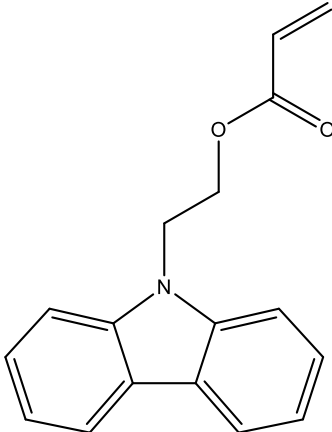
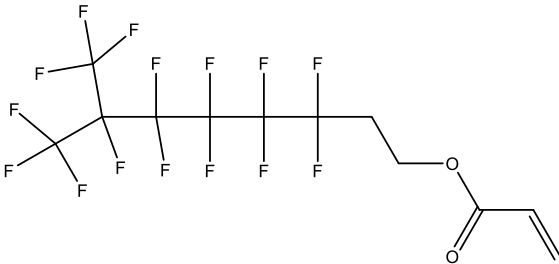
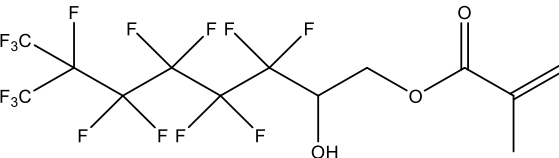
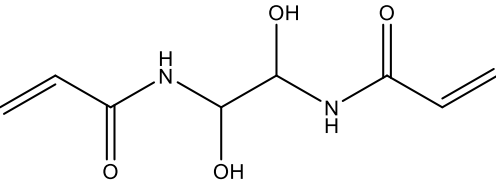
Acronym	Name	Structure
tBCHA	Tert-butylcyclohexylacrylate	
tBCHMA	Tertbutylcyclohexyl methacrylate	
tBMAm	N-tert-Butylmethacrylamide	
TEGDA	Tetra(ethylene glycol) diacrylate	
THFuA	Tetrahydrofurfuryl acrylate	
THFuMA	Tetrahydrofurfuryl methacrylate	

Acronym	Name	Structure
TMCHMA	Trimethylcyclohexyl methacrylate	
TMHA	Trimethylhexyl acrylate	
TMPDAE	Trimethyl propane diallyl ether	
TMPETA	Trimethylolpropane ethoxylate triacrylate	
TMPOTA	Trimethylolpropane propoxylate triacrylate	

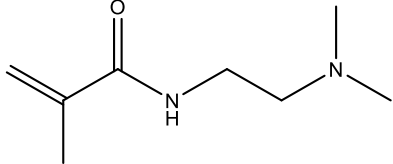
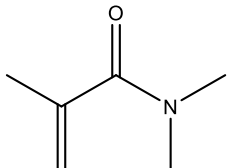
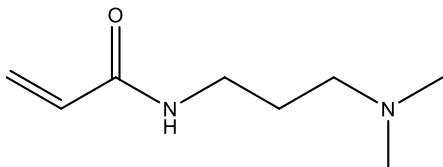
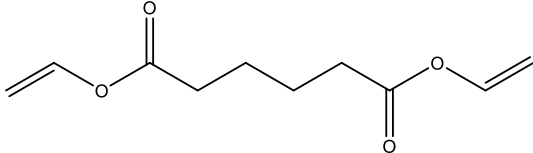
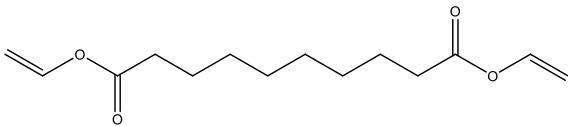
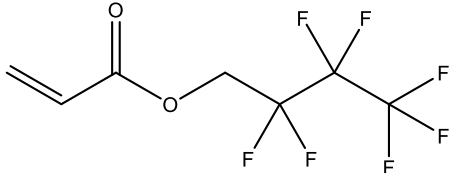
Acronym	Name	Structure
TPGDA	Tri(propylene glycol) diacrylate	
ZnA	Zinc acrylate	
ZrA	Zirconium acrylate	
ZrBNCTA	Zirconium bromonorbornanelactone carboxylate triacrylate	
AAM	Acrylamide	

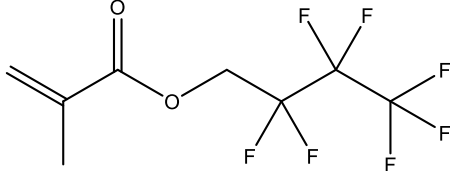
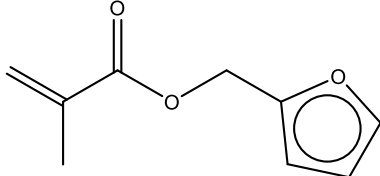
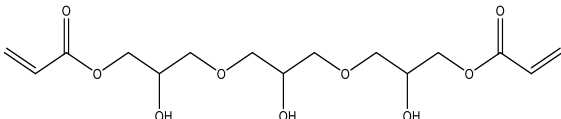
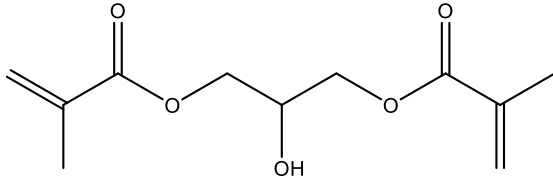
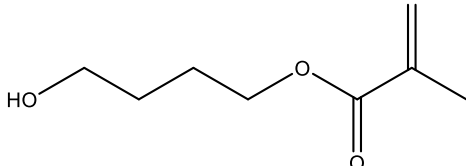
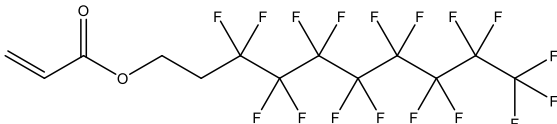
Acronym	Name	Structure
BAC	N,N'-Bis(acryloyl)cystamine	
BAPA	1,4-Bis(acryloyl)piperazine	
BAPODA	Bisphenol A propoxylate diacrylate	
BHMA	Benzhydryl methacrylate	
BMAOEP	Bis[2-(methacryloyloxy)ethyl] phosphate	

Acronym	Name	Structure
BMENBC	Bis(2-methacryloxyethyl) N,N'-1,9-nonylene biscarbamate	
BOEMA	Butoxyethyl methacrylate	
BPDMA	Bisphenol A dimethacrylate	
CEA	Carboxyethyl acrylate	
COEA	2-Cinnamoyloxyethyl acrylate	

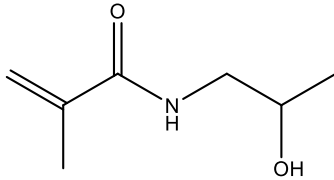
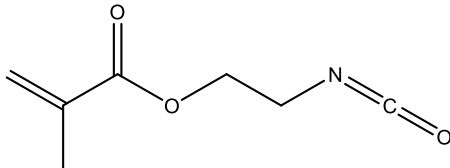
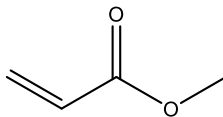
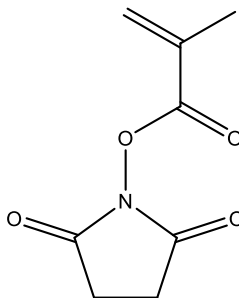
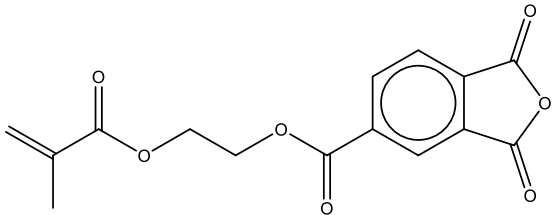
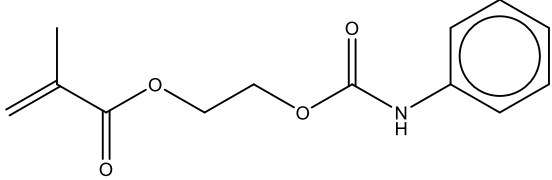
Acronym	Name	Structure
CzEA	Carbazol-9-yl ethyl acrylate	
DFFMOA	Dodecafluoro-7-(trifluoromethyl)-octyl acrylate	
DFHNMA	Dodecafluoro-2-hydroxy-8-(trifluoromethyl)nonyl methacrylate	
DHEBAM	N,N'-(1,2-Dihydroxyethylene)bisacrylamide	

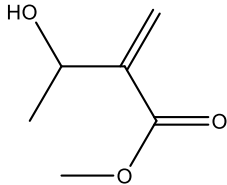
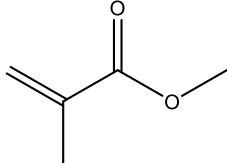
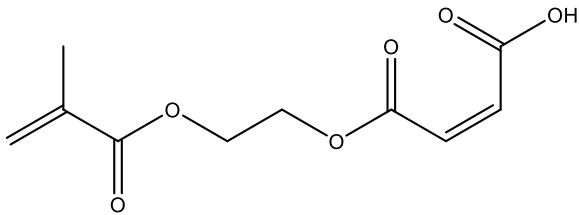
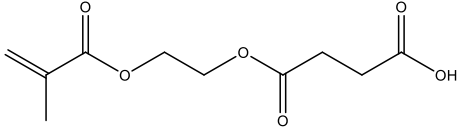
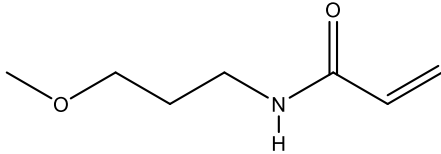
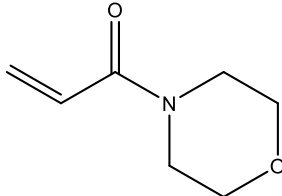


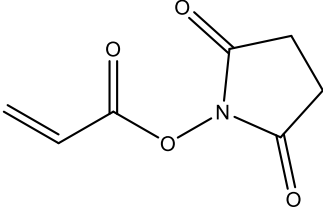
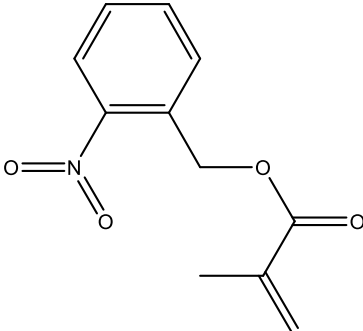
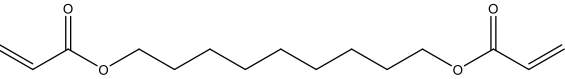
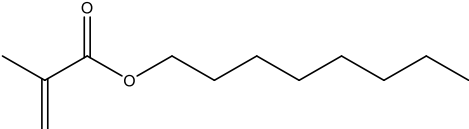
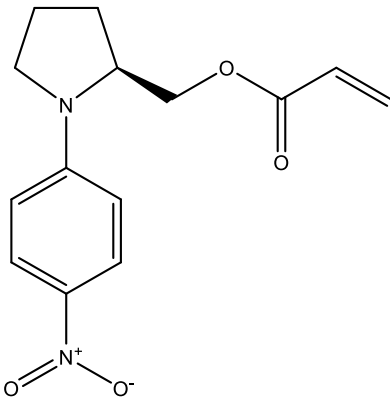
Acronym	Name	Structure
DMEMAm	N-[2-(N,N-Dimethylamino)ethyl]methacrylamide	
DMMAm	N,N-Dimethylmethacrylamide	
DMPAm	N-[3-(Dimethylamino)propyl]acrylamide	
DVAd	Divinyl Adipate	
DVSeb	Divinyl sebacate	
F7BA	Heptafluorobutyl acrylate	

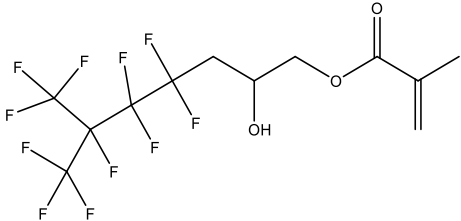
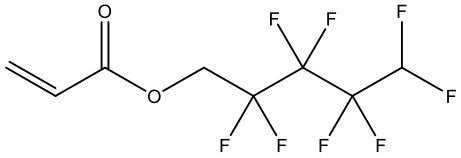
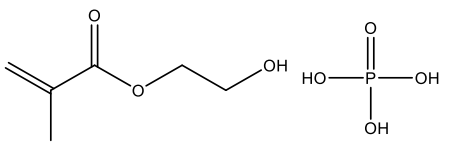
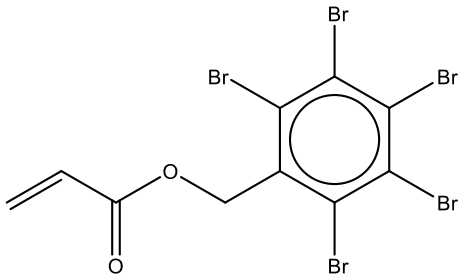
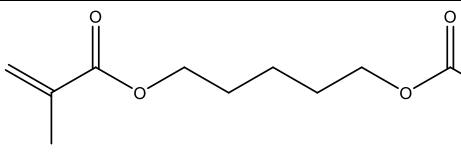
Acronym	Name	Structure
F7BMA	Heptafluorobutyl methacrylate	
FuMA	Furfuryl methacrylate	
GDGDA	Glycerol 1,3-diglycerolate diacrylate	
GDMA	Glycerol dimethacrylate	
HBMA	Hydroxybutyl methacrylate	
HDFDA	Hepta-decafluorodecyl acrylate	

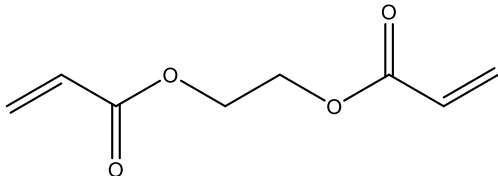
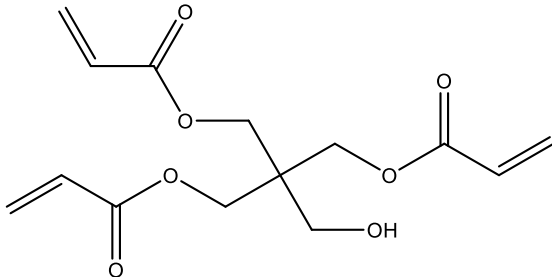
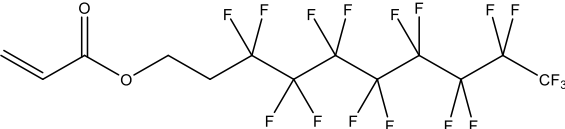
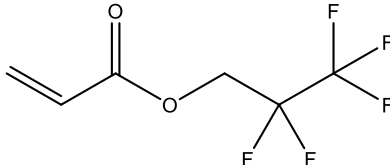
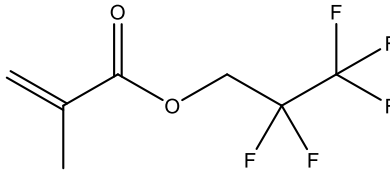
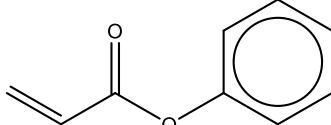
Acronym	Name	Structure
HDFDMA	Heptadecafluorodecyl methacrylate	
HDMA	1-Hexadecyl methacrylate	
HEODA	Hexanediol ethoxylate diacrylate	
HFDA	Heneicosafuorododecyl acrylate	
HFHUMA	Hexadecafluoro-2-hydroxy-10-(trifluoromethyl)undecyl methacrylate	
HMAm	N-(Hydroxymethyl)acrylamide	
HMBMAm	N,N'-Hexamethylenebis(methacrylamide)	

Acronym	Name	Structure
HPMAm	N-(2-Hydroxypropyl)methacrylamide	
iCEMA	Isocyanatoethyl methacrylate	
MA	Methyl acrylate	
MAAHS	Methacrylic acid N-hydroxysuccinimide ester	
MAETA	4-Methacryloxyethyl trimellitic anhydride	
MAPU	2-methacryloxyethyl phenyl urethane	

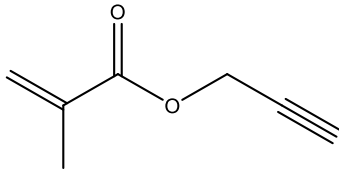
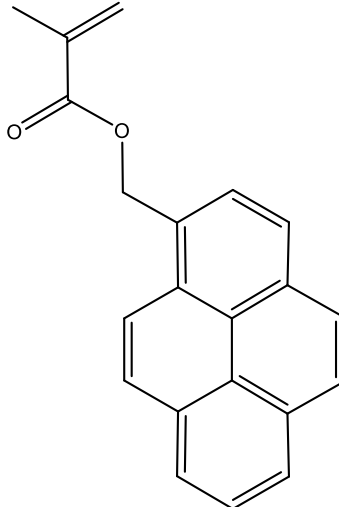
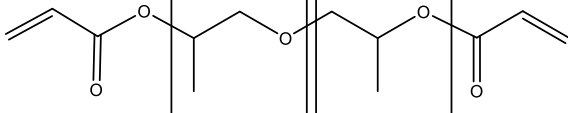
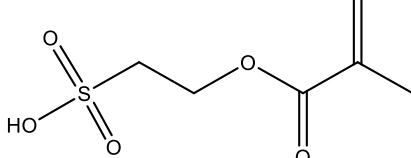
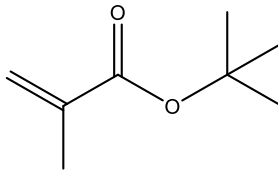
Acronym	Name	Structure
MHMB	Methyl 3-hydroxy-2-methylenebutyrate	
MMA	Methyl methacrylate	
mMAOEM	mono-2-(Methacryloyloxy)ethyl maleate	
mMAOES	mono-2-(Methacryloyloxy)ethyl succinate	
MOPAm	<i>N</i> -(3-Methoxypropyl)acrylamide	
NAM	<i>N</i> -Acryloylmorpholine	

Acronym	Name	Structure
NAS	<i>N</i> -Acryloxysuccinimide	
NBnMA	<i>o</i> -Nitrobenzyl methacrylate	
NDDMA	1,9-Nonanediol dimethacrylate	
nOcMA	<i>n</i> -Octyl methacrylate	
NPhPMA	Nitrophenyl-2-pyrrolidonemethyl acrylate	<div>  <div>Chiral</div> </div>

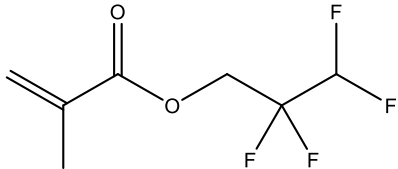
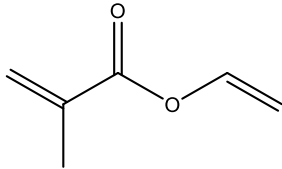
Acronym	Name	Structure
OFHMA	Octafluoro-2-hydroxy-6-(trifluoromethyl)heptyl methacrylate	
OFPA	Octafluoropentyl acrylate	
PAHEMA	Phosphoric acid 2-hydroxyethyl methacrylate ester	
PBBA	Pentabromobenzyl acrylate	
PDDMA	1,5-Pentanediol dimethacrylate	

Acronym	Name	Structure
pEGDA	Polyethylene glycol diacrylate	
PETrA	Pentaerythritol triacrylate	
pFDA	Perfluorodecyl acrylate	
PFPA	Pentafluoropropyl acrylate	
PFPMA	Pentafluoropropyl methacrylate	
PhA	Phenyl acrylate	



Acronym	Name	Structure
PMA	Propargyl methacrylate	
PMMA	1-Pyrenylmethyl methacrylate	
pPGDA	Poly(propylene glycol) diacrylate	
SEMA	2-Sulfoethyl methacrylate	
tBMA	Tert-butyl methacrylate	

Acronym	Name	Structure
TBNpMA	Tribromoneopentyl methacrylate	
TBPhA	2,4,6-Tribromophenyl acrylate	
TBPMA	Tribromophenyl methacrylate	
TCDMDA	Tricyclodecane-dimethanol diacrylate	
TDFOcA	Tridecafluorooctyl acrylate	
TEGMA	Tri(ethylene glycol) methyl ether methacrylate	

Acronym	Name	Structure
TFPMA	Tetrafluoropropyl methacrylate	
VMA	Vinyl methacrylate	
ZrCEA	Zirconium carboxyethyl acrylate	